

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Ward, Eric
Volrath, Sandra
Johnson, Marie
Potter, Sharon
- (ii) TITLE OF INVENTION: HERBICIDE TOLERANCE ACHIEVED THROUGH PLASTID TRANSFORMATION
- (iii) NUMBER OF SEQUENCES: 37
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Syngenta Biotechnology Inc.
 - (B) STREET: 3054 Cornwallis Road
 - (C) CITY: Research Triangle Park
 - (D) STATE: NC
 - (E) COUNTRY: USA
 - (F) ZIP: 27709
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 09/730,525
 - (B) FILING DATE: 5-DEC-2000
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 09/059,164
 - (B) FILING DATE: 13-APR-1998
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 09/050,603
 - (B) FILING DATE: 30-MAR-1998
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 09/038,878
 - (B) FILING DATE: 11-MAR-1998
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/808,931
 - (B) FILING DATE: 28-FEB-1997
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/012,705
 - (B) FILING DATE: 28-FEB-1996
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/013,612

(B) FILING DATE: 28-FEB-1996

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 60/020,003

(B) FILING DATE: 21-JUN-1996

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/472,028

(B) FILING DATE: 06-JUN-1995

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/261,198

(B) FILING DATE: 16-JUN-1994

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Kakefuda, Mary

(B) REGISTRATION NUMBER: 39,245

(C) REFERENCE/DOCKET NUMBER: 20757USCON8

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (919) 765-5071

(B) TELEFAX: (919) 541-8689

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

| | |
|--|-----|
| TGACAAAATT CCGAATTCTC TGC GATTTC ATG GAG TTA TCT CTT CTC CGT CCG | 54 |
| Met Glu Leu Ser Leu Leu Arg Pro | |
| 1 5 | |
| ACG ACT CAA TCG CTT CTT CCG TCG TTT TCG AAG CCC AAT CTC CGA TTA | 102 |
| Thr Thr Gln Ser Leu Leu Pro Ser Phe Ser Lys Pro Asn Leu Arg Leu | |
| 10 15 20 | |
| AAT GTT TAT AAG CCT CTT AGA CTC CGT TGT TCA GTG GCC GGT GGA CCA | 150 |
| Asn Val Tyr Lys Pro Leu Arg Leu Arg Cys Ser Val Ala Gly Gly Pro | |
| 25 30 35 40 | |
| ACC GTC GGA TCT TCA AAA ATC GAA GGC GGA GGA GGC ACC ACC ATC ACG | 198 |
| Thr Val Gly Ser Ser Lys Ile Glu Gly Gly Gly Thr Thr Ile Thr | |
| 45 50 55 | |
| ACG GAT TGT GTG ATT GTC GGC GGA GGT ATT AGT GGT CTT TGC ATC GCT | 246 |
| Thr Asp Cys Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala | |
| 60 65 70 | |
| CAG GCG CTT GCT ACT AAG CAT CCT GAT GCT GCT CCG AAT TTA ATT GTG | 294 |
| Gln Ala Leu Ala Thr Lys His Pro Asp Ala Ala Pro Asn Leu Ile Val | |
| 75 80 85 | |
| ACC GAG GCT AAG GAT CGT GTT GGA GGC AAC ATT ATC ACT CGT GAA GAG | 342 |
| Thr Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Ile Thr Arg Glu Glu | |
| 90 95 100 | |
| AAT GGT TTT CTC TGG GAA GAA GGT CCC AAT AGT TTT CAA CCG TCT GAT | 390 |
| Asn Gly Phe Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp | |
| 105 110 115 120 | |
| CCT ATG CTC ACT ATG GTG GTA GAT AGT GGT TTG AAG GAT GAT TTG GTG | 438 |
| Pro Met Leu Thr Met Val Val Asp Ser Gly Leu Lys Asp Asp Leu Val | |
| 125 130 135 | |

| | |
|---|------|
| TTG GGA GAT CCT ACT GCG CCA AGG TTT GTG TTG TGG AAT GGG AAA TTG Leu Gly Asp Pro Thr Ala Pro Arg Phe Val Leu Trp Asn Gly Lys Leu 140 145 150 | 486 |
| AGG CCG GTT CCA TCG AAG CTA ACA GAC TTA CCG TTC TTT GAT TTG ATG Arg Pro Val Pro Ser Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met 155 160 165 | 534 |
| AGT ATT GGT GGG AAG ATT AGA GCT GGT TTT GGT GCA CTT GGC ATT CGA Ser Ile Gly Gly Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg 170 175 180 | 582 |
| CCG TCA CCT CCA GGT CGT GAA GAA TCT GTG GAG GAG TTT GTA CGG CGT Pro Ser Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg 185 190 195 200 | 630 |
| AAC CTC GGT GAT GAG GTT TTT GAG CGC CTG ATT GAA CCG TTT TGT TCA Asn Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser 205 210 215 | 678 |
| GGT GTT TAT GCT GGT GAT CCT TCA AAA CTG AGC ATG AAA GCA GCG TTT Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe 220 225 230 | 726 |
| GGG AAG GTT TGG AAA CTA GAG CAA AAT GGT GGA AGC ATA ATA GGT GGT Gly Lys Val Trp Lys Leu Glu Gln Asn Gly Gly Ser Ile Ile Gly Gly 235 240 245 | 774 |
| ACT TTT AAG GCA ATT CAG GAG AGG AAA AAC GCT CCC AAG GCA GAA CGA Thr Phe Lys Ala Ile Gln Glu Arg Lys Asn Ala Pro Lys Ala Glu Arg 250 255 260 | 822 |
| GAC CCG CGC CTG CCA AAA CCA CAG GGC CAA ACA GTT GGT TCT TTC AGG Asp Pro Arg Leu Pro Lys Pro Gln Gly Gln Thr Val Gly Ser Phe Arg 265 270 275 280 | 870 |
| AAG GGA CTT CGA ATG TTG CCA GAA GCA ATA TCT GCA AGA TTA GGT AGC Lys Gly Leu Arg Met Leu Pro Glu Ala Ile Ser Ala Arg Leu Gly Ser 285 290 295 | 918 |
| AAA GTT AAG TTG TCT TGG AAG CTC TCA GGT ATC ACT AAG CTG GAG AGC Lys Val Lys Leu Ser Trp Lys Leu Ser Gly Ile Thr Lys Leu Glu Ser 300 305 310 | 966 |
| GGA GGA TAC AAC TTA ACA TAT GAG ACT CCA GAT GGT TTA GTT TCC GTG Gly Gly Tyr Asn Leu Thr Tyr Glu Thr Pro Asp Gly Leu Val Ser Val 315 320 325 | 1014 |
| CAG AGC AAA AGT GTT GTA ATG ACG GTG CCA TCT CAT GTT GCA AGT GGT Gln Ser Lys Ser Val Val Met Thr Val Pro Ser His Val Ala Ser Gly 330 335 340 | 1062 |
| CTC TTG CGC CCT CTT TCT GAA TCT GCT GCA AAT GCA CTC TCA AAA CTA Leu Leu Arg Pro Leu Ser Glu Ser Ala Ala Asn Ala Leu Ser Lys Leu 345 350 355 360 | 1110 |
| TAT TAC CCA CCA GTT GCA GCA GTA TCT ATC TCG TAC CCG AAA GAA GCA Tyr Tyr Pro Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala 365 370 375 | 1158 |
| ATC CGA ACA GAA TGT TTG ATA GAT GGT GAA CTA AAG GGT TTT GGG CAA | 1206 |

| | | | | | | | | | | | | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|------|
| Ile | Arg | Thr | Glu | Cys | Leu | Ile | Asp | Gly | Glu | Leu | Lys | Gly | Phe | Gly | Gln | | |
| | | | 380 | | | | | 385 | | | | | 390 | | | | |
| TTG | CAT | CCA | CGC | ACG | CAA | GGA | GTT | GAA | ACA | TTA | GGA | ACT | ATC | TAC | AGC | | 1254 |
| Leu | His | Pro | Arg | Thr | Gln | Gly | Val | Glu | Thr | Leu | Gly | Thr | Ile | Tyr | Ser | | |
| | | 395 | | | | 400 | | | | | 405 | | | | | | |
| TCC | TCA | CTC | TTT | CCA | AAT | CGC | GCA | CCG | CCC | GGA | AGA | ATT | TTG | CTG | TTG | | 1302 |
| Ser | Ser | Leu | Phe | Pro | Asn | Arg | Ala | Pro | Pro | Gly | Arg | Ile | Leu | Leu | Leu | | |
| | 410 | | | | 415 | | | | | 420 | | | | | | | |
| AAC | TAC | ATT | GGC | GGG | TCT | ACA | AAC | ACC | GGA | ATT | CTG | TCC | AAG | TCT | GAA | | 1350 |
| Asn | Tyr | Ile | Gly | Gly | Ser | Thr | Asn | Thr | Gly | Ile | Leu | Ser | Lys | Ser | Glu | | |
| 425 | | | | 430 | | | | | 435 | | | | | | 440 | | |
| GGT | GAG | TTA | GTG | GAA | GCA | GTT | GAC | AGA | GAT | TTG | AGG | AAA | ATG | CTA | ATT | | 1398 |
| Gly | Glu | Leu | Val | Glu | Ala | Val | Asp | Arg | Asp | Leu | Arg | Lys | Met | Leu | Ile | | |
| | | | 445 | | | | 450 | | | | | | 455 | | | | |
| AAG | CCT | AAT | TCG | ACC | GAT | CCA | CTT | AAA | TTA | GGA | GTT | AGG | GTA | TGG | CCT | | 1446 |
| Lys | Pro | Asn | Ser | Thr | Asp | Pro | Leu | Lys | Leu | Gly | Val | Arg | Val | Trp | Pro | | |
| | | 460 | | | | 465 | | | | | 470 | | | | | | |
| CAA | GCC | ATT | CCT | CAG | TTT | CTA | GTT | GGT | CAC | TTT | GAT | ATC | CTT | GAC | ACG | | 1494 |
| Gln | Ala | Ile | Pro | Gln | Phe | Leu | Val | Gly | His | Phe | Asp | Ile | Leu | Asp | Thr | | |
| | 475 | | | | | 480 | | | | | 485 | | | | | | |
| GCT | AAA | TCA | TCT | CTA | ACG | TCT | TCG | GGC | TAC | GAA | GGG | CTA | TTT | TTG | GGT | | 1542 |
| Ala | Lys | Ser | Ser | Leu | Thr | Ser | Ser | Gly | Tyr | Glu | Gly | Leu | Phe | Leu | Gly | | |
| | 490 | | | | 495 | | | | | 500 | | | | | | | |
| GGC | AAT | TAC | GTC | GCT | GGT | GTA | GCC | TTA | GGC | CGG | TGT | GTA | GAA | GGC | GCA | | 1590 |
| Gly | Asn | Tyr | Val | Ala | Gly | Val | Ala | Leu | Gly | Arg | Cys | Val | Glu | Gly | Ala | | |
| 505 | | | | 510 | | | | | 515 | | | | | | 520 | | |
| TAT | GAA | ACC | GCG | ATT | GAG | GTC | AAC | AAC | TTC | ATG | TCA | CGG | TAC | GCT | TAC | | 1638 |
| Tyr | Glu | Thr | Ala | Ile | Glu | Val | Asn | Asn | Phe | Met | Ser | Arg | Tyr | Ala | Tyr | | |
| | | | 525 | | | | 530 | | | | | | 535 | | | | |
| AAG | TAAATGTAAA | ACATTAAATC | TCCCAGCTTG | CGTGAGTTTT | ATTAAATATT | | | | | | | | | | | | 1691 |
| Lys | | | | | | | | | | | | | | | | | |
| TTGAGATATC | CAAAAAAAAA | AAAAAAAAA | | | | | | | | | | | | | | | 1719 |

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 537 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Glu | Leu | Ser | Leu | Leu | Arg | Pro | Thr | Thr | Gln | Ser | Leu | Leu | Pro | Ser |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |
| Phe | Ser | Lys | Pro | Asn | Leu | Arg | Leu | Asn | Val | Tyr | Lys | Pro | Leu | Arg | Leu |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | | 20 | | | | 25 | | | | 30 | | | | |
| Arg | Cys | Ser | Val | Ala | Gly | Gly | Pro | Thr | Val | Gly | Ser | Ser | Lys | Ile | Glu |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Gly | Gly | Gly | Gly | Thr | Thr | Ile | Thr | Thr | Asp | Cys | Val | Ile | Val | Gly | Gly |
| | | 50 | | | 55 | | | | | | 60 | | | | |
| Gly | Ile | Ser | Gly | Leu | Cys | Ile | Ala | Gln | Ala | Leu | Ala | Thr | Lys | His | Pro |
| | | 65 | | | 70 | | | | | 75 | | | 80 | | |
| Asp | Ala | Ala | Pro | Asn | Leu | Ile | Val | Thr | Glu | Ala | Lys | Asp | Arg | Val | Gly |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Gly | Asn | Ile | Ile | Thr | Arg | Glu | Glu | Asn | Gly | Phe | Leu | Trp | Glu | Glu | Gly |
| | | 100 | | | | 105 | | | | | | 110 | | | |
| Pro | Asn | Ser | Phe | Gln | Pro | Ser | Asp | Pro | Met | Leu | Thr | Met | Val | Val | Asp |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Ser | Gly | Leu | Lys | Asp | Asp | Leu | Val | Leu | Gly | Asp | Pro | Thr | Ala | Pro | Arg |
| | | 130 | | | 135 | | | | | | 140 | | | | |
| Phe | Val | Leu | Trp | Asn | Gly | Lys | Leu | Arg | Pro | Val | Pro | Ser | Lys | Leu | Thr |
| 145 | | | | 150 | | | | | | 155 | | | | | 160 |
| Asp | Leu | Pro | Phe | Phe | Asp | Leu | Met | Ser | Ile | Gly | Gly | Lys | Ile | Arg | Ala |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Gly | Phe | Gly | Ala | Leu | Gly | Ile | Arg | Pro | Ser | Pro | Pro | Gly | Arg | Glu | Glu |
| | | 180 | | | | 185 | | | | | | 190 | | | |
| Ser | Val | Glu | Glu | Phe | Val | Arg | Arg | Asn | Leu | Gly | Asp | Glu | Val | Phe | Glu |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Arg | Leu | Ile | Glu | Pro | Phe | Cys | Ser | Gly | Val | Tyr | Ala | Gly | Asp | Pro | Ser |
| | | 210 | | | 215 | | | | | | 220 | | | | |
| Lys | Leu | Ser | Met | Lys | Ala | Ala | Phe | Gly | Lys | Val | Trp | Lys | Leu | Glu | Gln |
| 225 | | | | 230 | | | | | | 235 | | | | | 240 |
| Asn | Gly | Gly | Ser | Ile | Ile | Gly | Gly | Thr | Phe | Lys | Ala | Ile | Gln | Glu | Arg |
| | | | | 245 | | | | | 250 | | | | | 255 | |
| Lys | Asn | Ala | Pro | Lys | Ala | Glu | Arg | Asp | Pro | Arg | Leu | Pro | Lys | Pro | Gln |
| | | 260 | | | | | | 265 | | | | | 270 | | |
| Gly | Gln | Thr | Val | Gly | Ser | Phe | Arg | Lys | Gly | Leu | Arg | Met | Leu | Pro | Glu |
| | | 275 | | | | 280 | | | | | | 285 | | | |
| Ala | Ile | Ser | Ala | Arg | Leu | Gly | Ser | Lys | Val | Lys | Leu | Ser | Trp | Lys | Leu |
| | | 290 | | | 295 | | | | | | 300 | | | | |
| Ser | Gly | Ile | Thr | Lys | Leu | Glu | Ser | Gly | Gly | Tyr | Asn | Leu | Thr | Tyr | Glu |
| 305 | | | | 310 | | | | | | 315 | | | | | 320 |
| Thr | Pro | Asp | Gly | Leu | Val | Ser | Val | Gln | Ser | Lys | Ser | Val | Val | Met | Thr |
| | | | | 325 | | | | | 330 | | | | | 335 | |
| Val | Pro | Ser | His | Val | Ala | Ser | Gly | Leu | Leu | Arg | Pro | Leu | Ser | Glu | Ser |
| | | 340 | | | | 345 | | | | | | 350 | | | |

Ala Ala Asn Ala Leu Ser Lys Leu Tyr Tyr Pro Pro Val Ala Ala Val
 355 360 365
 Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg Thr Glu Cys Leu Ile Asp
 370 375 380
 Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Thr Gln Gly Val
 385 390 395 400
 Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala
 405 410 415
 Pro Pro Gly Arg Ile Leu Leu Leu Asn Tyr Ile Gly Gly Ser Thr Asn
 420 425 430
 Thr Gly Ile Leu Ser Lys Ser Glu Gly Glu Leu Val Glu Ala Val Asp
 435 440 445
 Arg Asp Leu Arg Lys Met Leu Ile Lys Pro Asn Ser Thr Asp Pro Leu
 450 455 460
 Lys Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val
 465 470 475 480
 Gly His Phe Asp Ile Leu Asp Thr Ala Lys Ser Ser Leu Thr Ser Ser
 485 490 495
 Gly Tyr Glu Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala
 500 505 510
 Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala Ile Glu Val Asn
 515 520 525
 Asn Phe Met Ser Arg Tyr Ala Tyr Lys
 530 535

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1738 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Arabidopsis thaliana
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: pWDC-1 (NRRL B-21237)
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 70..1596

(D) OTHER INFORMATION: /product= "Arabidopsis protox-2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

| | |
|--|-----|
| TTTTTTACTT ATTTCCGTCA CTGCTTTCTGA CTGGTCAGAG ATTTTGACTC TGAATTGTTG | 60 |
| CAGATAGCA ATG GCG TCT GGA GCA GTA GCA GAT CAT CAA ATT GAA GCG | 108 |
| Met Ala Ser Gly Ala Val Ala Asp His Gln Ile Glu Ala | |
| 1 5 10 | |
| GTT TCA GGA AAA AGA GTC GCA GTC GTA GGT GCA GGT GTA AGT GGA CTT | 156 |
| Val Ser Gly Lys Arg Val Ala Val Val Gly Ala Gly Val Ser Gly Leu | |
| 15 20 25 | |
| GCG GCG GCT TAC AAG TTG AAA TCG AGG GGT TTG AAT GTG ACT GTG TTT | 204 |
| Ala Ala Ala Tyr Lys Leu Lys Ser Arg Gly Leu Asn Val Thr Val Phe | |
| 30 35 40 45 | |
| GAA GCT GAT GGA AGA GTA GGT GGG AAG TTG AGA AGT GTT ATG CAA AAT | 252 |
| Glu Ala Asp Gly Arg Val Gly Gly Lys Leu Arg Ser Val Met Gln Asn | |
| 50 55 60 | |
| GGT TTG ATT TGG GAT GAA GGA GCA AAC ACC ATG ACT GAG GCT GAG CCA | 300 |
| Gly Leu Ile Trp Asp Glu Gly Ala Asn Thr Met Thr Glu Ala Glu Pro | |
| 65 70 75 | |
| GAA GTT GGG AGT TTA CTT GAT GAT CTT GGG CTT CGT GAG AAA CAA CAA | 348 |
| Glu Val Gly Ser Leu Leu Asp Asp Leu Gly Leu Arg Glu Lys Gln Gln | |
| 80 85 90 | |
| TTT CCA ATT TCA CAG AAA AAG CGG TAT ATT GTG CGG AAT GGT GTA CCT | 396 |
| Phe Pro Ile Ser Gln Lys Lys Arg Tyr Ile Val Arg Asn Gly Val Pro | |
| 95 100 105 | |
| GTG ATG CTA CCT ACC AAT CCC ATA GAG CTG GTC ACA AGT AGT GTG CTC | 444 |
| Val Met Leu Pro Thr Asn Pro Ile Glu Leu Val Thr Ser Ser Val Leu | |
| 110 115 120 125 | |
| TCT ACC CAA TCT AAG TTT CAA ATC TTG TTG GAA CCA TTT TTA TGG AAG | 492 |
| Ser Thr Gln Ser Lys Phe Gln Ile Leu Leu Glu Pro Phe Leu Trp Lys | |
| 130 135 140 | |
| AAA AAG TCC TCA AAA GTC TCA GAT GCA TCT GCT GAA GAA AGT GTA AGC | 540 |
| Lys Lys Ser Ser Lys Val Ser Asp Ala Ser Ala Glu Glu Ser Val Ser | |
| 145 150 155 | |
| GAG TTC TTT CAA CGC CAT TTT GGA CAA GAG GTT GTT GAC TAT CTC ATC | 588 |
| Glu Phe Phe Gln Arg His Phe Gly Gln Glu Val Val Asp Tyr Leu Ile | |
| 160 165 170 | |
| GAC CCT TTT GTT GGT GGA ACA AGT GCT GCG GAC CCT GAT TCC CTT TCA | 636 |
| Asp Pro Phe Val Gly Gly Thr Ser Ala Ala Asp Pro Asp Ser Leu Ser | |
| 175 180 185 | |
| ATG AAG CAT TCT TTC CCA GAT CTC TGG AAT GTA GAG AAA AGT TTT GGC | 684 |
| Met Lys His Ser Phe Pro Asp Leu Trp Asn Val Glu Lys Ser Phe Gly | |
| 190 195 200 205 | |
| TCT ATT ATA GTC GGT GCA ATC AGA ACA AAG TTT GCT GCT AAA GGT GGT | 732 |
| Ser Ile Ile Val Gly Ala Ile Arg Thr Lys Phe Ala Ala Lys Gly Gly | |

| 210 | 215 | 220 | |
|---|-----|-----|------|
| AAA AGT AGA GAC ACA AAG AGT TCT CCT GGC ACA AAA AAG GGT TCG CGT Lys Ser Arg Asp Thr Lys Ser Ser Pro Gly Thr Lys Lys Gly Ser Arg 225 230 235 | | | 780 |
| GGG TCA TTC TCT TTT AAG GGG GGA ATG CAG ATT CTT CCT GAT ACG TTG Gly Ser Phe Ser Phe Lys Gly Gly Met Gln Ile Leu Pro Asp Thr Leu 240 245 250 | | | 828 |
| TGC AAA AGT CTC TCA CAT GAT GAG ATC AAT TTA GAC TCC AAG GTA CTC Cys Lys Ser Leu Ser His Asp Glu Ile Asn Leu Asp Ser Lys Val Leu 255 260 265 | | | 876 |
| TCT TTG TCT TAC AAT TCT GGA TCA AGA CAG GAG AAC TGG TCA TTA TCT Ser Leu Ser Tyr Asn Ser Gly Ser Arg Gln Glu Asn Trp Ser Leu Ser 270 275 280 285 | | | 924 |
| TGT GTT TCG CAT AAT GAA ACG CAG AGA CAA AAC CCC CAT TAT GAT GCT Cys Val Ser His Asn Glu Thr Gln Arg Gln Asn Pro His Tyr Asp Ala 290 295 300 | | | 972 |
| GTA ATT ATG ACG GCT CCT CTG TGC AAT GTG AAG GAG ATG AAG GTT ATG Val Ile Met Thr Ala Pro Leu Cys Asn Val Lys Glu Met Lys Val Met 305 310 315 | | | 1020 |
| AAA GGA GGA CAA CCC TTT CAG CTA AAC TTT CTC CCC GAG ATT AAT TAC Lys Gly Gly Gln Pro Phe Gln Leu Asn Phe Leu Pro Glu Ile Asn Tyr 320 325 330 | | | 1068 |
| ATG CCC CTC TCG GTT TTA ATC ACC ACA TTC ACA AAG GAG AAA GTA AAG Met Pro Leu Ser Val Leu Ile Thr Thr Phe Thr Lys Glu Lys Val Lys 335 340 345 | | | 1116 |
| AGA CCT CTT GAA GGC TTT GGG GTA CTC ATT CCA TCT AAG GAG CAA AAG Arg Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Ser Lys Glu Gln Lys 350 355 360 365 | | | 1164 |
| CAT GGT TTC AAA ACT CTA GGT ACA CTT TTT TCA TCA ATG ATG TTT CCA His Gly Phe Lys Thr Leu Gly Thr Leu Phe Ser Ser Met Met Phe Pro 370 375 380 | | | 1212 |
| GAT CGT TCC CCT AGT GAC GTT CAT CTA TAT ACA ACT TTT ATT GGT GGG Asp Arg Ser Pro Ser Asp Val His Leu Tyr Thr Thr Phe Ile Gly Gly 385 390 395 | | | 1260 |
| AGT AGG AAC CAG GAA CTA GCC AAA GCT TCC ACT GAC GAA TTA AAA CAA Ser Arg Asn Gln Glu Leu Ala Lys Ala Ser Thr Asp Glu Leu Lys Gln 400 405 410 | | | 1308 |
| GTT GTG ACT TCT GAC CTT CAG CGA CTG TTG GGG GTT GAA GGT GAA CCC Val Val Thr Ser Asp Leu Gln Arg Leu Leu Gly Val Glu Gly Glu Pro 415 420 425 | | | 1356 |
| GTG TCT GTC AAC CAT TAC TAT TGG AGG AAA GCA TTC CCG TTG TAT GAC Val Ser Val Asn His Tyr Tyr Trp Arg Lys Ala Phe Pro Leu Tyr Asp 430 435 440 445 | | | 1404 |
| AGC AGC TAT GAC TCA GTC ATG GAA GCA ATT GAC AAG ATG GAG AAT GAT Ser Ser Tyr Asp Ser Val Met Glu Ala Ile Asp Lys Met Glu Asn Asp 450 455 460 | | | 1452 |

CTA CCT GGG TTC TTC TAT GCA GGT AAT CAT CGA GGG GGG CTC TCT GTT 1500
 Leu Pro Gly Phe Phe Tyr Ala Gly Asn His Arg Gly Gly Leu Ser Val
 465 470 475

GGG AAA TCA ATA GCA TCA GGT TGC AAA GCA GCT GAC CTT GTG ATC TCA 1548
 Gly Lys Ser Ile Ala Ser Gly Cys Lys Ala Ala Asp Leu Val Ile Ser
 480 485 490

TAC CTG GAG TCT TGC TCA AAT GAC AAG AAA CCA AAT GAC AGC TTA TAACATTGTC 1603
 Tyr Leu Glu Ser Cys Ser Asn Asp Lys Lys Pro Asn Asp Ser Leu
 495 500 505

AAGGTTTCGTC CCTTTTATC ACTTACTTTG TAACTTGTA AAATGCAACA AGCCGCCGTG 1663

CGATTAGCCA ACAACTCAGC AAAACCCAGA TTCTCATAAG GCTCACTAAT TCCAGAATAA 1723

ACTATTTATG TAAAA 1738

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 508 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ala Ser Gly Ala Val Ala Asp His Gln Ile Glu Ala Val Ser Gly
 1 5 10 15

Lys Arg Val Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala
 20 25 30

Tyr Lys Leu Lys Ser Arg Gly Leu Asn Val Thr Val Phe Glu Ala Asp
 35 40 45

Gly Arg Val Gly Gly Lys Leu Arg Ser Val Met Gln Asn Gly Leu Ile
 50 55 60

Trp Asp Glu Gly Ala Asn Thr Met Thr Glu Ala Glu Pro Glu Val Gly
 65 70 75 80

Ser Leu Leu Asp Asp Leu Gly Leu Arg Glu Lys Gln Gln Phe Pro Ile
 85 90 95

Ser Gln Lys Lys Arg Tyr Ile Val Arg Asn Gly Val Pro Val Met Leu
 100 105 110

Pro Thr Asn Pro Ile Glu Leu Val Thr Ser Ser Val Leu Ser Thr Gln
 115 120 125

Ser Lys Phe Gln Ile Leu Leu Glu Pro Phe Leu Trp Lys Lys Lys Ser
 130 135 140

Ser Lys Val Ser Asp Ala Ser Ala Glu Glu Ser Val Ser Glu Phe Phe
 145 150 155 160

Gln Arg His Phe Gly Gln Glu Val Val Asp Tyr Leu Ile Asp Pro Phe
 165 170 175
 Val Gly Gly Thr Ser Ala Ala Asp Pro Asp Ser Leu Ser Met Lys His
 180 185 190
 Ser Phe Pro Asp Leu Trp Asn Val Glu Lys Ser Phe Gly Ser Ile Ile
 195 200 205
 Val Gly Ala Ile Arg Thr Lys Phe Ala Ala Lys Gly Gly Lys Ser Arg
 210 215 220
 Asp Thr Lys Ser Ser Pro Gly Thr Lys Lys Gly Ser Arg Gly Ser Phe
 225 230 235 240
 Ser Phe Lys Gly Gly Met Gln Ile Leu Pro Asp Thr Leu Cys Lys Ser
 245 250 255
 Leu Ser His Asp Glu Ile Asn Leu Asp Ser Lys Val Leu Ser Leu Ser
 260 265 270
 Tyr Asn Ser Gly Ser Arg Gln Glu Asn Trp Ser Leu Ser Cys Val Ser
 275 280 285
 His Asn Glu Thr Gln Arg Gln Asn Pro His Tyr Asp Ala Val Ile Met
 290 295 300
 Thr Ala Pro Leu Cys Asn Val Lys Glu Met Lys Val Met Lys Gly Gly
 305 310 315 320
 Gln Pro Phe Gln Leu Asn Phe Leu Pro Glu Ile Asn Tyr Met Pro Leu
 325 330 335
 Ser Val Leu Ile Thr Thr Phe Thr Lys Glu Lys Val Lys Arg Pro Leu
 340 345 350
 Glu Gly Phe Gly Val Leu Ile Pro Ser Lys Glu Gln Lys His Gly Phe
 355 360 365
 Lys Thr Leu Gly Thr Leu Phe Ser Ser Met Met Phe Pro Asp Arg Ser
 370 375 380
 Pro Ser Asp Val His Leu Tyr Thr Thr Phe Ile Gly Gly Ser Arg Asn
 385 390 395 400
 Gln Glu Leu Ala Lys Ala Ser Thr Asp Glu Leu Lys Gln Val Val Thr
 405 410 415
 Ser Asp Leu Gln Arg Leu Leu Gly Val Glu Gly Glu Pro Val Ser Val
 420 425 430
 Asn His Tyr Tyr Trp Arg Lys Ala Phe Pro Leu Tyr Asp Ser Ser Tyr
 435 440 445
 Asp Ser Val Met Glu Ala Ile Asp Lys Met Glu Asn Asp Leu Pro Gly
 450 455 460
 Phe Phe Tyr Ala Gly Asn His Arg Gly Gly Leu Ser Val Gly Lys Ser
 465 470 475 480
 Ile Ala Ser Gly Cys Lys Ala Ala Asp Leu Val Ile Ser Tyr Leu Glu

485

490

495

Ser Cys Ser Asn Asp Lys Lys Pro Asn Asp Ser Leu
 500 505

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1691 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Zea mays (maize)

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pWDC-4 (NRRL B-21260)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1443
- (D) OTHER INFORMATION: /product= "Maize protox-1

cDNA "

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

| | |
|---|-----|
| GCG GAC TGC GTC GTG GTG GGC GGA GGC ATC AGT GGC CTC TGC ACC GCG | 48 |
| Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu Cys Thr Ala | |
| 1 5 10 15 | |
| CAG GCG CTG GCC ACG CGG CAC GGC GTC GGG GAC GTG CTT GTC ACG GAG | 96 |
| Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu Val Thr Glu | |
| 20 25 30 | |
| GCC CGC GCC CGC CCC GGC GGC AAC ATT ACC ACC GTC GAG CGC CCC GAG | 144 |
| Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Glu | |
| 35 40 45 | |
| GAA GGG TAC CTC TGG GAG GAG GGT CCC AAC AGC TTC CAG CCC TCC GAC | 192 |
| Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp | |
| 50 55 60 | |
| CCC GTT CTC ACC ATG GCC GTG GAC AGC GGA CTG AAG GAT GAC TTG GTT | 240 |
| Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val | |
| 65 70 75 80 | |
| TTT GGG GAC CCA AAC GCG CCG CGT TTC GTG CTG TGG GAG GGG AAG CTG | 288 |
| Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu | |
| 85 90 95 | |
| AGG CCC GTG CCA TCC AAG CCC GCC GAC CTC CCG TTC TTC GAT CTC ATG | 336 |
| Arg Pro Val Pro Ser Lys Pro Ala Asp Leu Pro Phe Phe Asp Leu Met | |

| 100 | 105 | 110 | |
|---|-----|-----|------|
| AGC ATC CCA GGG AAG CTC AGG GCC GGT CTA GGC GCG CTT GGC ATC CGC Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg 115 120 125 | | | 384 |
| CCG CCT CCT CCA GGC CGC GAA GAG TCA GTG GAG GAG TTC GTG CGC CGC Pro Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg 130 135 140 | | | 432 |
| AAC CTC GGT GCT GAG GTC TTT GAG CGC CTC ATT GAG CCT TTC TGC TCA Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser 145 150 155 160 | | | 480 |
| GGT GTC TAT GCT GGT GAT CCT TCT AAG CTC AGC ATG AAG GCT GCA TTT Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe 165 170 175 | | | 528 |
| GGG AAG GTT TGG CGG TTG GAA GAA ACT GGA GGT AGT ATT ATT GGT GGA Gly Lys Val Trp Arg Leu Glu Glu Thr Gly Gly Ser Ile Ile Gly Gly 180 185 190 | | | 576 |
| ACC ATC AAG ACA ATT CAG GAG AGG AGC AAG AAT CCA AAA CCA CCG AGG Thr Ile Lys Thr Ile Gln Glu Arg Ser Lys Asn Pro Lys Pro Pro Arg 195 200 205 | | | 624 |
| GAT GCC CGC CTT CCG AAG CCA AAA GGG CAG ACA GTT GCA TCT TTC AGG Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Ala Ser Phe Arg 210 215 220 | | | 672 |
| AAG GGT CTT GCC ATG CTT CCA AAT GCC ATT ACA TCC AGC TTG GGT AGT Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Thr Ser Ser Leu Gly Ser 225 230 235 240 | | | 720 |
| AAA GTC AAA CTA TCA TGG AAA CTC ACG AGC ATT ACA AAA TCA GAT GAC Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ser Asp Asp 245 250 255 | | | 768 |
| AAG GGA TAT GTT TTG GAG TAT GAA ACG CCA GAA GGG GTT GTT TCG GTG Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro Glu Gly Val Val Ser Val 260 265 270 | | | 816 |
| CAG GCT AAA AGT GTT ATC ATG ACT ATT CCA TCA TAT GTT GCT AGC AAC Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asn 275 280 285 | | | 864 |
| ATT TTG CGT CCA CTT TCA AGC GAT GCT GCA GAT GCT CTA TCA AGA TTC Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala Asp Ala Leu Ser Arg Phe 290 295 300 | | | 912 |
| TAT TAT CCA CCG GTT GCT GCT GTA ACT GTT TCG TAT CCA AAG GAA GCA Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala 305 310 315 320 | | | 960 |
| ATT AGA AAA GAA TGC TTA ATT GAT GGG GAA CTC CAG GGC TTT GGC CAG Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln 325 330 335 | | | 1008 |
| TTG CAT CCA CGT AGT CAA GGA GTT GAG ACA TTA GGA ACA ATA TAC AGT Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser 340 345 350 | | | 1056 |

| | |
|---|------|
| TCC TCA CTC TTT CCA AAT CGT GCT CCT GAC GGT AGG GTG TTA CTT CTA | 1104 |
| Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val Leu Leu Leu | |
| 355 360 365 | |
| AAC TAC ATA GGA GGT GCT ACA AAC ACA GGA ATT GTT TCC AAG ACT GAA | 1152 |
| Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser Lys Thr Glu | |
| 370 375 380 | |
| AGT GAG CTG GTC GAA GCA GTT GAC CGT GAC CTC CGA AAA ATG CTT ATA | 1200 |
| Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile | |
| 385 390 395 400 | |
| AAT TCT ACA GCA GTG GAC CCT TTA GTC CTT GGT GTT CGA GTT TGG CCA | 1248 |
| Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg Val Trp Pro | |
| 405 410 415 | |
| CAA GCC ATA CCT CAG TTC CTG GTA GGA CAT CTT GAT CTT CTG GAA GCC | 1296 |
| Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Glu Ala | |
| 420 425 430 | |
| GCA AAA GCT GCC CTG GAC CGA GGT GGC TAC GAT GGG CTG TTC CTA GGA | 1344 |
| Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu Phe Leu Gly | |
| 435 440 445 | |
| GGG AAC TAT GTT GCA GGA GTT GCC CTG GGC AGA TGC GTT GAG GGC GCG | 1392 |
| Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala | |
| 450 455 460 | |
| TAT GAA AGT GCC TCG CAA ATA TCT GAC TTC TTG ACC AAG TAT GCC TAC | 1440 |
| Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys Tyr Ala Tyr | |
| 465 470 475 480 | |
| AAG TGATGAAAGA AGTGGAGCGC TACTTGTTAA TCGTTTATGT TGCATAGATG | 1493 |
| Lys | |
| AGGTGCCTCC GGGGAAAAAA AAGCTTGAAT AGTATTTTTT ATTCTTATTT TGTAATTGTC | 1553 |
| ATTTCTGTTC TTTTTTCTAT CAGTAATTAG TTATATTTTA GTTCTGTAGG AGATTGTTCT | 1613 |
| GTTCACTGCC CTTCAAAGA AATTTTATTT TTCATTCTTT TATGAGAGCT GTGCTACTTA | 1673 |
| AAAAAAAAA AAAAAAAAA | 1691 |

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 481 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

| |
|---|
| Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu Cys Thr Ala |
| 1 5 10 15 |
| Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu Val Thr Glu |

| 20 | | | | | 25 | | | | | 30 | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Arg | Ala | Arg | Pro | Gly | Gly | Asn | Ile | Thr | Thr | Val | Glu | Arg | Pro | Glu |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Glu | Gly | Tyr | Leu | Trp | Glu | Glu | Gly | Pro | Asn | Ser | Phe | Gln | Pro | Ser | Asp |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Pro | Val | Leu | Thr | Met | Ala | Val | Asp | Ser | Gly | Leu | Lys | Asp | Asp | Leu | Val |
| | 65 | | | | | 70 | | | | | 75 | | | | 80 |
| Phe | Gly | Asp | Pro | Asn | Ala | Pro | Arg | Phe | Val | Leu | Trp | Glu | Gly | Lys | Leu |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Arg | Pro | Val | Pro | Ser | Lys | Pro | Ala | Asp | Leu | Pro | Phe | Phe | Asp | Leu | Met |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Ser | Ile | Pro | Gly | Lys | Leu | Arg | Ala | Gly | Leu | Gly | Ala | Leu | Gly | Ile | Arg |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Pro | Pro | Pro | Pro | Gly | Arg | Glu | Glu | Ser | Val | Glu | Glu | Phe | Val | Arg | Arg |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Asn | Leu | Gly | Ala | Glu | Val | Phe | Glu | Arg | Leu | Ile | Glu | Pro | Phe | Cys | Ser |
| | 145 | | | | | 150 | | | | | 155 | | | | 160 |
| Gly | Val | Tyr | Ala | Gly | Asp | Pro | Ser | Lys | Leu | Ser | Met | Lys | Ala | Ala | Phe |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Gly | Lys | Val | Trp | Arg | Leu | Glu | Glu | Thr | Gly | Gly | Ser | Ile | Ile | Gly | Gly |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Thr | Ile | Lys | Thr | Ile | Gln | Glu | Arg | Ser | Lys | Asn | Pro | Lys | Pro | Pro | Arg |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Asp | Ala | Arg | Leu | Pro | Lys | Pro | Lys | Gly | Gln | Thr | Val | Ala | Ser | Phe | Arg |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Lys | Gly | Leu | Ala | Met | Leu | Pro | Asn | Ala | Ile | Thr | Ser | Ser | Leu | Gly | Ser |
| | 225 | | | | | 230 | | | | | 235 | | | 240 | |
| Lys | Val | Lys | Leu | Ser | Trp | Lys | Leu | Thr | Ser | Ile | Thr | Lys | Ser | Asp | Asp |
| | | | 245 | | | | | 250 | | | | | | 255 | |
| Lys | Gly | Tyr | Val | Leu | Glu | Tyr | Glu | Thr | Pro | Glu | Gly | Val | Val | Ser | Val |
| | | | 260 | | | | 265 | | | | | | 270 | | |
| Gln | Ala | Lys | Ser | Val | Ile | Met | Thr | Ile | Pro | Ser | Tyr | Val | Ala | Ser | Asn |
| | | 275 | | | | | 280 | | | | | 285 | | | |
| Ile | Leu | Arg | Pro | Leu | Ser | Ser | Asp | Ala | Ala | Asp | Ala | Leu | Ser | Arg | Phe |
| | 290 | | | | | 295 | | | | | 300 | | | | |
| Tyr | Tyr | Pro | Pro | Val | Ala | Ala | Val | Thr | Val | Ser | Tyr | Pro | Lys | Glu | Ala |
| | 305 | | | | | 310 | | | | | 315 | | | 320 | |
| Ile | Arg | Lys | Glu | Cys | Leu | Ile | Asp | Gly | Glu | Leu | Gln | Gly | Phe | Gly | Gln |
| | | | 325 | | | | | | 330 | | | | | 335 | |
| Leu | His | Pro | Arg | Ser | Gln | Gly | Val | Glu | Thr | Leu | Gly | Thr | Ile | Tyr | Ser |
| | | | 340 | | | | | 345 | | | | | 350 | | |

Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val Leu Leu Leu
 355 360 365

Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser Lys Thr Glu
 370 375 380

Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile
 385 390 395 400

Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg Val Trp Pro
 405 410 415

Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Glu Ala
 420 425 430

Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu Phe Leu Gly
 435 440 445

Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala
 450 455 460

Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys Tyr Ala Tyr
 465 470 475 480

Lys

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2061 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Zea mays (maize)
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: pWDC-3 (NRRL B-21259)
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 64..1698
 - (D) OTHER INFORMATION: /product= "Maize protox-2"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTCTCCTACC TCCACCTCCA CGACAACAAG CAAATCCCCA TCCAGTTCCA AACCCCTAACT 60

CAA ATG CTC GCT TTG ACT GCC TCA GCC TCA TCC GCT TCG TCC CAT CCT 108

Met Leu Ala Leu Thr Ala Ser Ala Ser Ser Ala Ser Ser His Pro

1 5 10 15

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| TAT | CGC | CAC | GCC | TCC | GCG | CAC | ACT | CGT | CGC | CCC | CGC | CTA | CGT | GCG | GTC | 156 |
| Tyr | Arg | His | Ala | Ser | Ala | His | Thr | Arg | Arg | Pro | Arg | Leu | Arg | Ala | Val | |
| | | | 20 | | | | | | 25 | | | | | 30 | | |
| CTC | GCG | ATG | GCG | GGC | TCC | GAC | GAC | CCC | CGT | GCA | GCG | CCC | GCC | AGA | TCG | 204 |
| Leu | Ala | Met | Ala | Gly | Ser | Asp | Asp | Pro | Arg | Ala | Ala | Pro | Ala | Arg | Ser | |
| | | | 35 | | | | | 40 | | | | | 45 | | | |
| GTC | GCC | GTC | GTC | GGC | GCC | GGG | GTC | AGC | GGG | CTC | GCG | GCG | GCG | TAC | AGG | 252 |
| Val | Ala | Val | Val | Gly | Ala | Gly | Val | Ser | Gly | Leu | Ala | Ala | Ala | Tyr | Arg | |
| | | | 50 | | | | 55 | | | | | | 60 | | | |
| CTC | AGA | CAG | AGC | GGC | GTG | AAC | GTA | ACG | GTG | TTC | GAA | GCG | GCC | GAC | AGG | 300 |
| Leu | Arg | Gln | Ser | Gly | Val | Asn | Val | Thr | Val | Phe | Glu | Ala | Ala | Asp | Arg | |
| | 65 | | | | | 70 | | | | | 75 | | | | | |
| GCG | GGA | GGA | AAG | ATA | CGG | ACC | AAT | TCC | GAG | GGC | GGG | TTT | GTC | TGG | GAT | 348 |
| Ala | Gly | Gly | Lys | Ile | Arg | Thr | Asn | Ser | Glu | Gly | Gly | Phe | Val | Trp | Asp | |
| | 80 | | | | 85 | | | | | 90 | | | | | 95 | |
| GAA | GGA | GCT | AAC | ACC | ATG | ACA | GAA | GGT | GAA | TGG | GAG | GCC | AGT | AGA | CTG | 396 |
| Glu | Gly | Ala | Asn | Thr | Met | Thr | Glu | Gly | Glu | Trp | Glu | Ala | Ser | Arg | Leu | |
| | | | | 100 | | | | | 105 | | | | | 110 | | |
| ATT | GAT | GAT | CTT | GGT | CTA | CAA | GAC | AAA | CAG | CAG | TAT | CCT | AAC | TCC | CAA | 444 |
| Ile | Asp | Asp | Leu | Gly | Leu | Gln | Asp | Lys | Gln | Gln | Tyr | Pro | Asn | Ser | Gln | |
| | | | 115 | | | | | 120 | | | | | 125 | | | |
| CAC | AAG | CGT | TAC | ATT | GTC | AAA | GAT | GGA | GCA | CCA | GCA | CTG | ATT | CCT | TCG | 492 |
| His | Lys | Arg | Tyr | Ile | Val | Lys | Asp | Gly | Ala | Pro | Ala | Leu | Ile | Pro | Ser | |
| | | 130 | | | | | 135 | | | | | 140 | | | | |
| GAT | CCC | ATT | TCG | CTA | ATG | AAA | AGC | AGT | GTT | CTT | TCG | ACA | AAA | TCA | AAG | 540 |
| Asp | Pro | Ile | Ser | Leu | Met | Lys | Ser | Ser | Val | Leu | Ser | Thr | Lys | Ser | Lys | |
| | 145 | | | | | 150 | | | | | 155 | | | | | |
| ATT | GCG | TTA | TTT | TTT | GAA | CCA | TTT | CTC | TAC | AAG | AAA | GCT | AAC | ACA | AGA | 588 |
| Ile | Ala | Leu | Phe | Phe | Glu | Pro | Phe | Leu | Tyr | Lys | Lys | Ala | Asn | Thr | Arg | |
| | 160 | | | | 165 | | | | | 170 | | | | | 175 | |
| AAC | TCT | GGA | AAA | GTG | TCT | GAG | GAG | CAC | TTG | AGT | GAG | AGT | GTT | GGG | AGC | 636 |
| Asn | Ser | Gly | Lys | Val | Ser | Glu | Glu | His | Leu | Ser | Glu | Ser | Val | Gly | Ser | |
| | | | | 180 | | | | | 185 | | | | | 190 | | |
| TTC | TGT | GAA | CGC | CAC | TTT | GGA | AGA | GAA | GTT | GTT | GAC | TAT | TTT | GTT | GAT | 684 |
| Phe | Cys | Glu | Arg | His | Phe | Gly | Arg | Glu | Val | Val | Asp | Tyr | Phe | Val | Asp | |
| | | | 195 | | | | | 200 | | | | | 205 | | | |
| CCA | TTT | GTA | GCT | GGA | ACA | AGT | GCA | GGA | GAT | CCA | GAG | TCA | CTA | TCT | ATT | 732 |
| Pro | Phe | Val | Ala | Gly | Thr | Ser | Ala | Gly | Asp | Pro | Glu | Ser | Leu | Ser | Ile | |
| | | 210 | | | | | 215 | | | | | 220 | | | | |
| CGT | CAT | GCA | TTC | CCA | GCA | TTG | TGG | AAT | TTG | GAA | AGA | AAG | TAT | GGT | TCA | 780 |
| Arg | His | Ala | Phe | Pro | Ala | Leu | Trp | Asn | Leu | Glu | Arg | Lys | Tyr | Gly | Ser | |
| | 225 | | | | | 230 | | | | | 235 | | | | | |
| GTT | ATT | GTT | GGT | GCC | ATC | TTG | TCT | AAG | CTA | GCA | GCT | AAA | GGT | GAT | CCA | 828 |
| Val | Ile | Val | Gly | Ala | Ile | Leu | Ser | Lys | Leu | Ala | Ala | Lys | Gly | Asp | Pro | |
| | 240 | | | | 245 | | | | | 250 | | | | 255 | | |

| | |
|---|------|
| GTA AAG ACA AGA CAT GAT TCA TCA GGG AAA AGA AGG AAT AGA CGA GTG Val Lys Thr Arg His Asp Ser Ser Gly Lys Arg Arg Asn Arg Arg Val 260 265 270 | 876 |
| TCG TTT TCA TTT CAT GGT GGA ATG CAG TCA CTA ATA AAT GCA CTT CAC Ser Phe Ser Phe His Gly Gly Met Gln Ser Leu Ile Asn Ala Leu His 275 280 285 | 924 |
| AAT GAA GTT GGA GAT GAT AAT GTG AAG CTT GGT ACA GAA GTG TTG TCA Asn Glu Val Gly Asp Asp Asn Val Lys Leu Gly Thr Glu Val Leu Ser 290 295 300 | 972 |
| TTG GCA TGT ACA TTT GAT GGA GTT CCT GCA CTA GGC AGG TGG TCA ATT Leu Ala Cys Thr Phe Asp Gly Val Pro Ala Leu Gly Arg Trp Ser Ile 305 310 315 | 1020 |
| TCT GTT GAT TCG AAG GAT AGC GGT GAC AAG GAC CTT GCT AGT AAC CAA Ser Val Asp Ser Lys Asp Ser Gly Asp Lys Asp Leu Ala Ser Asn Gln 320 325 330 335 | 1068 |
| ACC TTT GAT GCT GTT ATA ATG ACA GCT CCA TTG TCA AAT GTC CGG AGG Thr Phe Asp Ala Val Ile Met Thr Ala Pro Leu Ser Asn Val Arg Arg 340 345 350 | 1116 |
| ATG AAG TTC ACC AAA GGT GGA GCT CCG GTT GTT CTT GAC TTT CTT CCT Met Lys Phe Thr Lys Gly Gly Ala Pro Val Val Leu Asp Phe Leu Pro 355 360 365 | 1164 |
| AAG ATG GAT TAT CTA CCA CTA TCT CTC ATG GTG ACT GCT TTT AAG AAG Lys Met Asp Tyr Leu Pro Leu Ser Leu Met Val Thr Ala Phe Lys Lys 370 375 380 | 1212 |
| GAT GAT GTC AAG AAA CCT CTG GAA GGA TTT GGG GTC TTA ATA CCT TAC Asp Asp Val Lys Lys Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Tyr 385 390 395 | 1260 |
| AAG GAA CAG CAA AAA CAT GGT CTG AAA ACC CTT GGG ACT CTC TTT TCC Lys Glu Gln Gln Lys His Gly Leu Lys Thr Leu Gly Thr Leu Phe Ser 400 405 410 415 | 1308 |
| TCA ATG ATG TTC CCA GAT CGA GCT CCT GAT GAC CAA TAT TTA TAT ACA Ser Met Met Phe Pro Asp Arg Ala Pro Asp Asp Gln Tyr Leu Tyr Thr 420 425 430 | 1356 |
| ACA TTT GTT GGG GGT AGC CAC AAT AGA GAT CTT GCT GGA GCT CCA ACG Thr Phe Val Gly Gly Ser His Asn Arg Asp Leu Ala Gly Ala Pro Thr 435 440 445 | 1404 |
| TCT ATT CTG AAA CAA CTT GTG ACC TCT GAC CTT AAA AAA CTC TTG GGC Ser Ile Leu Lys Gln Leu Val Thr Ser Asp Leu Lys Lys Leu Leu Gly 450 455 460 | 1452 |
| GTA GAG GGG CAA CCA ACT TTT GTC AAG CAT GTA TAC TGG GGA AAT GCT Val Glu Gly Gln Pro Thr Phe Val Lys His Val Tyr Trp Gly Asn Ala 465 470 475 | 1500 |
| TTT CCT TTG TAT GGC CAT GAT TAT AGT TCT GTA TTG GAA GCT ATA GAA Phe Pro Leu Tyr Gly His Asp Tyr Ser Ser Val Leu Glu Ala Ile Glu 480 485 490 495 | 1548 |
| AAG ATG GAG AAA AAC CTT CCA GGG TTC TTC TAC GCA GGA AAT AGC AAG | 1596 |

Lys Met Glu Lys Asn Leu Pro Gly Phe Phe Tyr Ala Gly Asn Ser Lys
 500 505 510
 GAT GGG CTT GCT GTT GGA AGT GTT ATA GCT TCA GGA AGC AAG GCT GCT 1644
 Asp Gly Leu Ala Val Gly Ser Val Ile Ala Ser Gly Ser Lys Ala Ala
 515 520 525
 GAC CTT GCA ATC TCA TAT CTT GAA TCT CAC ACC AAG CAT AAT AAT TCA 1692
 Asp Leu Ala Ile Ser Tyr Leu Glu Ser His Thr Lys His Asn Asn Ser
 530 535 540
 CAT TGAAAGTGTC TGACCTATCC TCTAGCAGTT GTCGACAAAT TTCTCCAGTT 1745
 His
 545
 CATGTACAGT AGAAACCGAT GCGTTGCAGT TTCAGAACAT CTTCACTTCT TCAGATATTA 1805
 ACCCTTCGTT GAACATCCAC CAGAAAGGTA GTCACATGTG TAAGTGGGAA AATGAGGTTA 1865
 AAAACTATTA TGGCGGCCGA AATGTTTCCTT TTTGTTTTCC TCACAAGTGG CCTACGACAC 1925
 TTGATGTTGG AAATACATTT AAATTTGTTG AATTGTTTGA GAACACATGC GTGACGTGTA 1985
 ATATTTGCCT ATTGTGATTT TAGCAGTAGT CTTGGCCAGA TTATGCTTTA CGCCTTTAAA 2045
 AAAAAAAAAA AAAAAA 2061

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 544 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Leu Ala Leu Thr Ala Ser Ala Ser Ser Ala Ser Ser His Pro Tyr
 1 5 10 15
 Arg His Ala Ser Ala His Thr Arg Arg Pro Arg Leu Arg Ala Val Leu
 20 25 30
 Ala Met Ala Gly Ser Asp Asp Pro Arg Ala Ala Pro Ala Arg Ser Val
 35 40 45
 Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala Tyr Arg Leu
 50 55 60
 Arg Gln Ser Gly Val Asn Val Thr Val Phe Glu Ala Ala Asp Arg Ala
 65 70 75 80
 Gly Gly Lys Ile Arg Thr Asn Ser Glu Gly Gly Phe Val Trp Asp Glu
 85 90 95
 Gly Ala Asn Thr Met Thr Glu Gly Glu Trp Glu Ala Ser Arg Leu Ile
 100 105 110
 Asp Asp Leu Gly Leu Gln Asp Lys Gln Gln Tyr Pro Asn Ser Gln His

| 115 | | | | | 120 | | | | | 125 | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Arg | Tyr | Ile | Val | Lys | Asp | Gly | Ala | Pro | Ala | Leu | Ile | Pro | Ser | Asp |
| 130 | | | | | 135 | | | | | 140 | | | | | |
| Pro | Ile | Ser | Leu | Met | Lys | Ser | Ser | Val | Leu | Ser | Thr | Lys | Ser | Lys | Ile |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Ala | Leu | Phe | Phe | Glu | Pro | Phe | Leu | Tyr | Lys | Lys | Ala | Asn | Thr | Arg | Asn |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Ser | Gly | Lys | Val | Ser | Glu | Glu | His | Leu | Ser | Glu | Ser | Val | Gly | Ser | Phe |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Cys | Glu | Arg | His | Phe | Gly | Arg | Glu | Val | Val | Asp | Tyr | Phe | Val | Asp | Pro |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Phe | Val | Ala | Gly | Thr | Ser | Ala | Gly | Asp | Pro | Glu | Ser | Leu | Ser | Ile | Arg |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| His | Ala | Phe | Pro | Ala | Leu | Trp | Asn | Leu | Glu | Arg | Lys | Tyr | Gly | Ser | Val |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| Ile | Val | Gly | Ala | Ile | Leu | Ser | Lys | Leu | Ala | Ala | Lys | Gly | Asp | Pro | Val |
| | | | 245 | | | | | | 250 | | | | | 255 | |
| Lys | Thr | Arg | His | Asp | Ser | Ser | Gly | Lys | Arg | Arg | Asn | Arg | Arg | Val | Ser |
| | | | 260 | | | | | 265 | | | | | 270 | | |
| Phe | Ser | Phe | His | Gly | Gly | Met | Gln | Ser | Leu | Ile | Asn | Ala | Leu | His | Asn |
| | | 275 | | | | | 280 | | | | | 285 | | | |
| Glu | Val | Gly | Asp | Asp | Asn | Val | Lys | Leu | Gly | Thr | Glu | Val | Leu | Ser | Leu |
| | 290 | | | | | 295 | | | | | 300 | | | | |
| Ala | Cys | Thr | Phe | Asp | Gly | Val | Pro | Ala | Leu | Gly | Arg | Trp | Ser | Ile | Ser |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 |
| Val | Asp | Ser | Lys | Asp | Ser | Gly | Asp | Lys | Asp | Leu | Ala | Ser | Asn | Gln | Thr |
| | | | | 325 | | | | | 330 | | | | | 335 | |
| Phe | Asp | Ala | Val | Ile | Met | Thr | Ala | Pro | Leu | Ser | Asn | Val | Arg | Arg | Met |
| | | | 340 | | | | | 345 | | | | | 350 | | |
| Lys | Phe | Thr | Lys | Gly | Gly | Ala | Pro | Val | Val | Leu | Asp | Phe | Leu | Pro | Lys |
| | | 355 | | | | | 360 | | | | | 365 | | | |
| Met | Asp | Tyr | Leu | Pro | Leu | Ser | Leu | Met | Val | Thr | Ala | Phe | Lys | Lys | Asp |
| | 370 | | | | | 375 | | | | | 380 | | | | |
| Asp | Val | Lys | Lys | Pro | Leu | Glu | Gly | Phe | Gly | Val | Leu | Ile | Pro | Tyr | Lys |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 |
| Glu | Gln | Gln | Lys | His | Gly | Leu | Lys | Thr | Leu | Gly | Thr | Leu | Phe | Ser | Ser |
| | | | | 405 | | | | | 410 | | | | | 415 | |
| Met | Met | Phe | Pro | Asp | Arg | Ala | Pro | Asp | Asp | Gln | Tyr | Leu | Tyr | Thr | Thr |
| | | | 420 | | | | | 425 | | | | | 430 | | |
| Phe | Val | Gly | Gly | Ser | His | Asn | Arg | Asp | Leu | Ala | Gly | Ala | Pro | Thr | Ser |
| | | 435 | | | | | 440 | | | | | 445 | | | |

```

Ile Leu Lys Gln Leu Val Thr Ser Asp Leu Lys Lys Leu Leu Gly Val
  450                      455                      460

Glu Gly Gln Pro Thr Phe Val Lys His Val Tyr Trp Gly Asn Ala Phe
465                      470                      475                      480

Pro Leu Tyr Gly His Asp Tyr Ser Ser Val Leu Glu Ala Ile Glu Lys
                      485                      490                      495

Met Glu Lys Asn Leu Pro Gly Phe Phe Tyr Ala Gly Asn Ser Lys Asp
                      500                      505                      510

Gly Leu Ala Val Gly Ser Val Ile Ala Ser Gly Ser Lys Ala Ala Asp
  515                      520                      525

Leu Ala Ile Ser Tyr Leu Glu Ser His Thr Lys His Asn Asn Ser His
  530                      535                      540

```

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1811 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Triticum aestivum (wheat)
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: pWDC-13 (NRRL B-21545)
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 3..1589
 - (D) OTHER INFORMATION: /product= "wheat protox-1"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

GC GCA ACA ATG GCC ACC GCC ACC GTC GCG GCC GCG TCG CCG CTC CGC      47
Ala Thr Met Ala Thr Ala Thr Val Ala Ala Ala Ser Pro Leu Arg
  1                      5                      10                      15

GGC AGG GTC ACC GGG CGC CCA CAC CGC GTC CGC CCG CGT TGC GCT ACC      95
Gly Arg Val Thr Gly Arg Pro His Arg Val Arg Pro Arg Cys Ala Thr
                      20                      25                      30

GCG AGC AGC GCG ACC GAG ACT CCG GCG GCG CCC GGC GTG CGG CTG TCC     143
Ala Ser Ser Ala Thr Glu Thr Pro Ala Ala Pro Gly Val Arg Leu Ser
                      35                      40                      45

GCG GAA TGC GTC ATT GTG GGC GCC GGC ATC AGC GGC CTC TGC ACC GCG     191
Ala Glu Cys Val Ile Val Gly Ala Gly Ile Ser Gly Leu Cys Thr Ala

```

| 50 | 55 | 60 | |
|---|----|----|-----|
| CAG GCG CTG GCC ACC CGA TAC GGC GTC AGC GAC CTG CTC GTC ACG GAG Gln Ala Leu Ala Thr Arg Tyr Gly Val Ser Asp Leu Leu Val Thr Glu 65 70 75 | | | 239 |
| GCC CGC GAC CGC CCG GGC GGC AAC ATC ACC ACC GTC GAG CGT CCC GAC Ala Arg Asp Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Asp 80 85 90 95 | | | 287 |
| GAG GGG TAC CTG TGG GAG GAG GGA CCC AAC AGC TTC CAG CCC TCC GAC Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp 100 105 110 | | | 335 |
| CCG GTC CTC ACC ATG GCC GTG GAC AGC GGG CTC AAG GAT GAC TTG GTG Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val 115 120 125 | | | 383 |
| TTC GGG GAC CCC AAC GCG CCC CGG TTC GTG CTG TGG GAG GGG AAG CTG Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu 130 135 140 | | | 431 |
| AGG CCG GTG CCG TCG AAG CCA GGC GAC CTG CCT TTC TTC AGC CTC ATG Arg Pro Val Pro Ser Lys Pro Gly Asp Leu Pro Phe Phe Ser Leu Met 145 150 155 | | | 479 |
| AGT ATC CCT GGG AAG CTC AGG GCC GGC CTT GGC GCG CTC GGC ATT CGC Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg 160 165 170 175 | | | 527 |
| CCA CCT CCT CCA GGG CGC GAG GAG TCG GTG GAG GAG TTT GTG CGC CGC Pro Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg 180 185 190 | | | 575 |
| AAC CTC GGT GCC GAG GTC TTT GAG CGC CTC ATC GAG CCT TTC TGC TCA Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser 195 200 205 | | | 623 |
| GGT GTA TAT GCT GGT GAT CCT TCG AAG CTT AGT ATG AAG GCT GCA TTT Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe 210 215 220 | | | 671 |
| GGG AAG GTC TGG AGG TTG GAG GAG ATT GGA GGT AGT ATT ATT GGT GGA Gly Lys Val Trp Arg Leu Glu Glu Ile Gly Gly Ser Ile Ile Gly Gly 225 230 235 | | | 719 |
| ACC ATC AAG GCG ATT CAG GAT AAA GGG AAG AAC CCC AAA CCG CCA AGG Thr Ile Lys Ala Ile Gln Asp Lys Gly Lys Asn Pro Lys Pro Pro Arg 240 245 250 255 | | | 767 |
| GAT CCC CGA CTT CCG GCA CCA AAG GGA CAG ACG GTG GCA TCT TTC AGG Asp Pro Arg Leu Pro Ala Pro Lys Gly Gln Thr Val Ala Ser Phe Arg 260 265 270 | | | 815 |
| AAG GGT CTA GCC ATG CTC CCG AAT GCC ATC GCA TCT AGG CTG GGT AGT Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Ala Ser Arg Leu Gly Ser 275 280 285 | | | 863 |
| AAA GTC AAG CTG TCA TGG AAG CTT ACG AGC ATT ACA AAG GCG GAC AAC Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ala Asp Asn 290 295 300 | | | 911 |

| | |
|---|------|
| CAA GGA TAT GTA TTA GGT TAT GAA ACA CCA GAA GGA CTT GTT TCA GTG Gln Gly Tyr Val Leu Gly Tyr Glu Thr Pro Glu Gly Leu Val Ser Val 305 310 315 | 959 |
| CAG GCT AAA AGT GTT ATC ATG ACC ATC CCG TCA TAT GTT GCT AGT GAT Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asp 320 325 330 335 | 1007 |
| ATC TTG CGC CCA CTT TCA ATT GAT GCA GCA GAT GCA CTC TCA AAA TTC Ile Leu Arg Pro Leu Ser Ile Asp Ala Ala Asp Ala Leu Ser Lys Phe 340 345 350 | 1055 |
| TAT TAT CCG CCA GTT GCT GCT GTA ACT GTT TCA TAT CCA AAA GAA GCT Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala 355 360 365 | 1103 |
| ATT AGA AAA GAA TGC TTA ATT GAT GGG GAG CTC CAG GGT TTC GGC CAG Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln 370 375 380 | 1151 |
| TTG CAT CCA CGT AGC CAA GGA GTC GAG ACT TTA GGG ACA ATA TAT AGC Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser 385 390 395 | 1199 |
| TCT TCT CTC TTT CCT AAT CGT GCT CCT GCT GGA AGA GTG TTA CTT CTG Ser Ser Leu Phe Pro Asn Arg Ala Pro Ala Gly Arg Val Leu Leu Leu 400 405 410 415 | 1247 |
| AAC TAT ATC GGG GGT TCT ACA AAT ACA GGG ATC GTC TCC AAG ACT GAG Asn Tyr Ile Gly Gly Ser Thr Asn Thr Gly Ile Val Ser Lys Thr Glu 420 425 430 | 1295 |
| AGT GAC TTA GTA GGA GCC GTT GAC CGT GAC CTC AGA AAA ATG TTG ATA Ser Asp Leu Val Gly Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile 435 440 445 | 1343 |
| AAC CCT AGA GCA GCA GAC CCT TTA GCA TTA GGG GTT CGA GTG TGG CCA Asn Pro Arg Ala Ala Asp Pro Leu Ala Leu Gly Val Arg Val Trp Pro 450 455 460 | 1391 |
| CAA GCA ATA CCA CAG TTT TTG ATT GGG CAC CTT GAT CGC CTT GCT GCT Gln Ala Ile Pro Gln Phe Leu Ile Gly His Leu Asp Arg Leu Ala Ala 465 470 475 | 1439 |
| GCA AAA TCT GCA CTG GGC CAA GGC GGC TAC GAC GGG TTG TTC CTA GGA Ala Lys Ser Ala Leu Gly Gln Gly Gly Tyr Asp Gly Leu Phe Leu Gly 480 485 490 495 | 1487 |
| GGA AAC TAC GTC GCA GGA GTT GCC TTG GGC CGA TGC ATC GAG GGT GCG Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala 500 505 510 | 1535 |
| TAC GAG AGT GCC TCA CAA GTA TCT GAC TTC TTG ACC AAG TAT GCC TAC Tyr Glu Ser Ala Ser Gln Val Ser Asp Phe Leu Thr Lys Tyr Ala Tyr 515 520 525 | 1583 |
| AAG TGA TGGAAGTAGT GCATCTCTTC ATTTTGTTC ATATACGAGG TGAGGCTAGG Lys | 1639 |

ATCGGTAAAA CATCATGAGA TTCTGTAGTG TTTCTTTAAT TGAAAAACA AATTTTAGTG 1699
 ATGCAATATG TGCTCTTTCC TGTAATTCGA GCATGTACAT CGGTATGGGA TAAAGTAGAA 1759
 TAAGCTATTC TGCAAAAGCA GTGATTTTTT TTGAAAAAAA AAAAAAAAAA AA 1811

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 528 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ala Thr Met Ala Thr Ala Thr Val Ala Ala Ala Ser Pro Leu Arg Gly
 1 5 10 15
 Arg Val Thr Gly Arg Pro His Arg Val Arg Pro Arg Cys Ala Thr Ala
 20 25 30
 Ser Ser Ala Thr Glu Thr Pro Ala Ala Pro Gly Val Arg Leu Ser Ala
 35 40 45
 Glu Cys Val Ile Val Gly Ala Gly Ile Ser Gly Leu Cys Thr Ala Gln
 50 55 60
 Ala Leu Ala Thr Arg Tyr Gly Val Ser Asp Leu Leu Val Thr Glu Ala
 65 70 75 80
 Arg Asp Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Asp Glu
 85 90 95
 Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro
 100 105 110
 Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val Phe
 115 120 125
 Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu Arg
 130 135 140
 Pro Val Pro Ser Lys Pro Gly Asp Leu Pro Phe Phe Ser Leu Met Ser
 145 150 155 160
 Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg Pro
 165 170 175
 Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn
 180 185 190
 Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly
 195 200 205
 Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly
 210 215 220
 Lys Val Trp Arg Leu Glu Glu Ile Gly Gly Ser Ile Ile Gly Gly Thr

| | | | | | | |
|---|--|-----|--|-----|--|-----|
| 225 | | 230 | | 235 | | 240 |
| Ile Lys Ala Ile Gln Asp Lys Gly Lys Asn Pro Lys Pro Pro Arg Asp | | | | | | |
| | | 245 | | 250 | | 255 |
| Pro Arg Leu Pro Ala Pro Lys Gly Gln Thr Val Ala Ser Phe Arg Lys | | | | | | |
| | | 260 | | 265 | | 270 |
| Gly Leu Ala Met Leu Pro Asn Ala Ile Ala Ser Arg Leu Gly Ser Lys | | | | | | |
| | | 275 | | 280 | | 285 |
| Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ala Asp Asn Gln | | | | | | |
| | | 290 | | 295 | | 300 |
| Gly Tyr Val Leu Gly Tyr Glu Thr Pro Glu Gly Leu Val Ser Val Gln | | | | | | |
| 305 | | 310 | | 315 | | 320 |
| Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asp Ile | | | | | | |
| | | 325 | | 330 | | 335 |
| Leu Arg Pro Leu Ser Ile Asp Ala Ala Asp Ala Leu Ser Lys Phe Tyr | | | | | | |
| | | 340 | | 345 | | 350 |
| Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala Ile | | | | | | |
| | | 355 | | 360 | | 365 |
| Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln Leu | | | | | | |
| | | 370 | | 375 | | 380 |
| His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser | | | | | | |
| 385 | | 390 | | 395 | | 400 |
| Ser Leu Phe Pro Asn Arg Ala Pro Ala Gly Arg Val Leu Leu Leu Asn | | | | | | |
| | | 405 | | 410 | | 415 |
| Tyr Ile Gly Gly Ser Thr Asn Thr Gly Ile Val Ser Lys Thr Glu Ser | | | | | | |
| | | 420 | | 425 | | 430 |
| Asp Leu Val Gly Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile Asn | | | | | | |
| | | 435 | | 440 | | 445 |
| Pro Arg Ala Ala Asp Pro Leu Ala Leu Gly Val Arg Val Trp Pro Gln | | | | | | |
| | | 450 | | 455 | | 460 |
| Ala Ile Pro Gln Phe Leu Ile Gly His Leu Asp Arg Leu Ala Ala Ala | | | | | | |
| 465 | | 470 | | 475 | | 480 |
| Lys Ser Ala Leu Gly Gln Gly Gly Tyr Asp Gly Leu Phe Leu Gly Gly | | | | | | |
| | | 485 | | 490 | | 495 |
| Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala Tyr | | | | | | |
| | | 500 | | 505 | | 510 |
| Glu Ser Ala Ser Gln Val Ser Asp Phe Leu Thr Lys Tyr Ala Tyr Lys | | | | | | |
| | | 515 | | 520 | | 525 |

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1847 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: soybean

(vii) IMMEDIATE SOURCE:

(B) CLONE: pWDC-12 (NRRL B-21516)

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 55..1683

(D) OTHER INFORMATION: /product= "soybean protox-1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

| | |
|---|-----|
| CTTTAGCACA GTGTTGAAGA TAACGAACGA ATAGTGCCAT TACTGTAACC AACC ATG | 57 |
| Met | |
| 1 | |
| GTT TCC GTC TTC AAC GAG ATC CTA TTC CCG CCG AAC CAA ACC CTT CTT | 105 |
| Val Ser Val Phe Asn Glu Ile Leu Phe Pro Pro Asn Gln Thr Leu Leu | |
| 5 10 15 | |
| CGC CCC TCC CTC CAT TCC CCA ACC TCT TTC TTC ACC TCT CCC ACT CGA | 153 |
| Arg Pro Ser Leu His Ser Pro Thr Ser Phe Phe Thr Ser Pro Thr Arg | |
| 20 25 30 | |
| AAA TTC CCT CGC TCT CGC CCT AAC CCT ATT CTA CGC TGC TCC ATT GCG | 201 |
| Lys Phe Pro Arg Ser Arg Pro Asn Pro Ile Leu Arg Cys Ser Ile Ala | |
| 35 40 45 | |
| GAG GAA TCC ACC GCG TCT CCG CCC AAA ACC AGA GAC TCC GCC CCC GTG | 249 |
| Glu Glu Ser Thr Ala Ser Pro Pro Lys Thr Arg Asp Ser Ala Pro Val | |
| 50 55 60 65 | |
| GAC TGC GTC GTC GTC GGC GGA GGC GTC AGC GGC CTC TGC ATC GCC CAG | 297 |
| Asp Cys Val Val Gly Gly Gly Val Ser Gly Leu Cys Ile Ala Gln | |
| 70 75 80 | |
| GCC CTC GCC ACC AAA CAC GCC AAT GCC AAC GTC GTC GTC ACG GAG GCC | 345 |
| Ala Leu Ala Thr Lys His Ala Asn Ala Asn Val Val Val Thr Glu Ala | |
| 85 90 95 | |
| CGA GAC CGC GTC GGC GGC AAC ATC ACC ACG ATG GAG AGG GAC GGA TAC | 393 |
| Arg Asp Arg Val Gly Gly Asn Ile Thr Thr Met Glu Arg Asp Gly Tyr | |
| 100 105 110 | |
| CTC TGG GAA GAA GGC CCC AAC AGC TTC CAG CCT TCT GAT CCA ATG CTC | 441 |
| Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu | |
| 115 120 125 | |
| ACC ATG GTG GTG GAC AGT GGT TTA AAG GAT GAG CTT GTT TTG GGG GAT | 489 |
| Thr Met Val Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu Gly Asp | |
| 130 135 140 145 | |

| | |
|---|------|
| CCT GAT GCA CCT CGG TTT GTG TTG TGG AAC AGG AAG TTG AGG CCG GTG Pro Asp Ala Pro Arg Phe Val Leu Trp Asn Arg Lys Leu Arg Pro Val 150 155 160 | 537 |
| CCC GGG AAG CTG ACT GAT TTG CCT TTC TTT GAC TTG ATG AGC ATT GGT Pro Gly Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly 165 170 175 | 585 |
| GGC AAA ATC AGG GCT GGC TTT GGT GCG CTT GGA ATT CGG CCT CCT CCT Gly Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg Pro Pro Pro 180 185 190 | 633 |
| CCA GGT CAT GAG GAA TCG GTT GAA GAG TTT GTT CGT CGG AAC CTT GGT Pro Gly His Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly 195 200 205 | 681 |
| GAT GAG GTT TTT GAA CGG TTG ATA GAG CCT TTT TGT TCA GGG GTC TAT Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr 210 215 220 225 | 729 |
| GCA GGC GAT CCT TCA AAA TTA AGT ATG AAA GCA GCA TTC GGG AAA GTT Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Lys Val 230 235 240 | 777 |
| TGG AAG CTG GAA AAA AAT GGT GGT AGC ATT ATT GGT GGA ACT TTC AAA Trp Lys Leu Glu Lys Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys 245 250 255 | 825 |
| GCA ATA CAA GAG AGA AAT GGA GCT TCA AAA CCA CCT CGA GAT CCG CGT Ala Ile Gln Glu Arg Asn Gly Ala Ser Lys Pro Pro Arg Asp Pro Arg 260 265 270 | 873 |
| CTG CCA AAA CCA AAA GGT CAG ACT GTT GGA TCT TTC CGG AAG GGA CTT Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu 275 280 285 | 921 |
| ACC ATG TTG CCT GAT GCA ATT TCT GCC AGA CTA GGC AAC AAA GTA AAG Thr Met Leu Pro Asp Ala Ile Ser Ala Arg Leu Gly Asn Lys Val Lys 290 295 300 305 | 969 |
| TTA TCT TGG AAG CTT TCA AGT ATT AGT AAA CTG GAT AGT GGA GAG TAC Leu Ser Trp Lys Leu Ser Ser Ile Ser Lys Leu Asp Ser Gly Glu Tyr 310 315 320 | 1017 |
| AGT TTG ACA TAT GAA ACA CCA GAA GGA GTG GTT TCT TTG CAG TGC AAA Ser Leu Thr Tyr Glu Thr Pro Glu Gly Val Val Ser Leu Gln Cys Lys 325 330 335 | 1065 |
| ACT GTT GTC CTG ACC ATT CCT TCC TAT GTT GCT AGT ACA TTG CTG CGT Thr Val Val Leu Thr Ile Pro Ser Tyr Val Ala Ser Thr Leu Leu Arg 340 345 350 | 1113 |
| CCT CTG TCT GCT GCT GCT GCA GAT GCA CTT TCA AAG TTT TAT TAC CCT Pro Leu Ser Ala Ala Ala Asp Ala Leu Ser Lys Phe Tyr Tyr Pro 355 360 365 | 1161 |
| CCA GTT GCT GCA GTT TCC ATA TCC TAT CCA AAA GAA GCT ATT AGA TCA Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg Ser 370 375 380 385 | 1209 |

| | |
|---|------|
| GAA TGC TTG ATA GAT GGT GAG TTG AAG GGG TTT GGT CAA TTG CAT CCA Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro 390 395 400 | 1257 |
| CGT AGC CAA GGA GTG GAA ACA TTA GGA ACT ATA TAC AGC TCA TCA CTA Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu 405 410 415 | 1305 |
| TTC CCC AAC CGA GCA CCA CCT GGA AGG GTT CTA CTC TTG AAT TAC ATT Phe Pro Asn Arg Ala Pro Pro Gly Arg Val Leu Leu Leu Asn Tyr Ile 420 425 430 | 1353 |
| GGA GGA GCA ACT AAT ACT GGA ATT TTA TCG AAG ACG GAC AGT GAA CTT Gly Gly Ala Thr Asn Thr Gly Ile Leu Ser Lys Thr Asp Ser Glu Leu 435 440 445 | 1401 |
| GTG GAA ACA GTT GAT CGA GAT TTG AGG AAA ATC CTT ATA AAC CCA AAT Val Glu Thr Val Asp Arg Asp Leu Arg Lys Ile Leu Ile Asn Pro Asn 450 455 460 465 | 1449 |
| GCC CAG GAT CCA TTT GTA GTG GGG GTG AGA CTG TGG CCT CAA GCT ATT Ala Gln Asp Pro Phe Val Val Gly Val Arg Leu Trp Pro Gln Ala Ile 470 475 480 | 1497 |
| CCA CAG TTC TTA GTT GGC CAT CTT GAT CTT CTA GAT GTT GCT AAA GCT Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Asp Val Ala Lys Ala 485 490 495 | 1545 |
| TCT ATC AGA AAT ACT GGG TTT GAA GGG CTC TTC CTT GGG GGT AAT TAT Ser Ile Arg Asn Thr Gly Phe Glu Gly Leu Phe Leu Gly Gly Asn Tyr 500 505 510 | 1593 |
| GTG TCT GGT GTT GCC TTG GGA CGA TGC GTT GAG GGA GCC TAT GAG GTA Val Ser Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Val 515 520 525 | 1641 |
| GCA GCT GAA GTA AAC GAT TTT CTC ACA AAT AGA GTG TAC AAA Ala Ala Glu Val Asn Asp Phe Leu Thr Asn Arg Val Tyr Lys 530 535 540 | 1683 |
| TAGTAGCAGT TTTTGT TTTT GTGGTGGAAT GGGTGATGGG ACTCTCGTGT TCCATTGAAT 1743 | |
| TATAATAATG TGAAAGTTTC TCAAATTCGT TCGATAGGTT TTTGGCGGCT TCTATTGCTG 1803 | |
| ATAATGTAAA ATCCTCTTTA AGTTTGAAAA AAAAAAAAAA AAAA 1847 | |

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 543 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Val | Ser | Val | Phe | Asn | Glu | Ile | Leu | Phe | Pro | Pro | Asn | Gln | Thr | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |

Leu Arg Pro Ser Leu His Ser Pro Thr Ser Phe Phe Thr Ser Pro Thr
 20 25 30
 Arg Lys Phe Pro Arg Ser Arg Pro Asn Pro Ile Leu Arg Cys Ser Ile
 35 40 45
 Ala Glu Glu Ser Thr Ala Ser Pro Pro Lys Thr Arg Asp Ser Ala Pro
 50 55 60
 Val Asp Cys Val Val Val Gly Gly Gly Val Ser Gly Leu Cys Ile Ala
 65 70 75 80
 Gln Ala Leu Ala Thr Lys His Ala Asn Ala Asn Val Val Val Thr Glu
 85 90 95
 Ala Arg Asp Arg Val Gly Gly Asn Ile Thr Thr Met Glu Arg Asp Gly
 100 105 110
 Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Met
 115 120 125
 Leu Thr Met Val Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu Gly
 130 135 140
 Asp Pro Asp Ala Pro Arg Phe Val Leu Trp Asn Arg Lys Leu Arg Pro
 145 150 155 160
 Val Pro Gly Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile
 165 170 175
 Gly Gly Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg Pro Pro
 180 185 190
 Pro Pro Gly His Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu
 195 200 205
 Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly Val
 210 215 220
 Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Lys
 225 230 235 240
 Val Trp Lys Leu Glu Lys Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe
 245 250 255
 Lys Ala Ile Gln Glu Arg Asn Gly Ala Ser Lys Pro Pro Arg Asp Pro
 260 265 270
 Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly
 275 280 285
 Leu Thr Met Leu Pro Asp Ala Ile Ser Ala Arg Leu Gly Asn Lys Val
 290 295 300
 Lys Leu Ser Trp Lys Leu Ser Ser Ile Ser Lys Leu Asp Ser Gly Glu
 305 310 315 320
 Tyr Ser Leu Thr Tyr Glu Thr Pro Glu Gly Val Val Ser Leu Gln Cys
 325 330 335
 Lys Thr Val Val Leu Thr Ile Pro Ser Tyr Val Ala Ser Thr Leu Leu

| 340 | 345 | 350 |
|---|-----|-----|
| Arg Pro Leu Ser Ala Ala Ala Asp Ala Leu Ser Lys Phe Tyr Tyr | | |
| 355 | 360 | 365 |
| Pro Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg | | |
| 370 | 375 | 380 |
| Ser Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His | | |
| 385 | 390 | 395 |
| Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser | | |
| | 405 | 410 |
| Leu Phe Pro Asn Arg Ala Pro Pro Gly Arg Val Leu Leu Leu Asn Tyr | | |
| | 420 | 425 |
| Ile Gly Gly Ala Thr Asn Thr Gly Ile Leu Ser Lys Thr Asp Ser Glu | | |
| | 435 | 440 |
| Leu Val Glu Thr Val Asp Arg Asp Leu Arg Lys Ile Leu Ile Asn Pro | | |
| | 450 | 455 |
| Asn Ala Gln Asp Pro Phe Val Val Gly Val Arg Leu Trp Pro Gln Ala | | |
| 465 | 470 | 475 |
| Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Asp Val Ala Lys | | |
| | 485 | 490 |
| Ala Ser Ile Arg Asn Thr Gly Phe Glu Gly Leu Phe Leu Gly Gly Asn | | |
| | 500 | 505 |
| Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu | | |
| | 515 | 520 |
| Val Ala Ala Glu Val Asn Asp Phe Leu Thr Asn Arg Val Tyr Lys | | |
| | 530 | 535 |
| | | 540 |

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 583 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (ix) FEATURE:
 - (A) NAME/KEY: promoter
 - (B) LOCATION: 1..583
 - (D) OTHER INFORMATION: /function= "arabidopsis protox-1 promoter"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAATTCCGAT CGAATTATAT AATTATCATA AATTTGAATA AGCATGTTGC CTTTATTAA

60

| | |
|--|-----|
| AGAGGTTTAA TAAAGTTTGG TAATAATGGA CTTTGACTTC AAACCTCGATT CTCATGTAAT | 120 |
| TAATTAATAT TTACATCAAA ATTTGGTCAC TAATATTACC AAATTAATAT ACTAAAAATGT | 180 |
| TAATTCGCAA ATAAACACT AATTCCAAAT AAAGGGTCAT TATGATAAAC ACGTATTGAA | 240 |
| CTTGATAAAG CAAAGCAAAA ATAATGGGTT TCAAGGTTTG GGTTATATAT GACAAAAAAA | 300 |
| AAAAAAGGTT TGGTTATATA TCTATTGGGC CTATAACCAT GTTATACAAA TTTGGGCCTA | 360 |
| ACTAAAATAA TAAAATAAAC GTAATGGTCC TTTTATATT TGGGTCAAAC CCAACTCTAA | 420 |
| ACCCAAACCA AAGAAAAAGT ATACGGTACG GTACACAGAC TTATGGTGTG TGTGATTGCA | 480 |
| GGTGAATATT TCTCGTCGTC TTCTCCTTTC TTCTGAAGAA GATTACCCAA TCTGAAAAAA | 540 |
| ACCAAGAAGC TGACAAAATT CCGAATTCTC TGCGATTTC ATG | 583 |

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3848 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 1..3848
- (D) OTHER INFORMATION: /function= "maize protox-1 promoter"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

| | |
|--|-----|
| TCGATCTTTC TAGGCTGATC CCCAAATCTT CCTCCGAAGC CCCTGGCGCC TCTGCCCCCTT | 60 |
| GGAGCTGGTG GCCTGAAAGA GCTTTGCTGT TGCCCCGAAG ATTGTGAGGT ATATTGTGAC | 120 |
| CTCTGAGACT GACTTCCTTT GTCGTCACCT TGAGTGGAGT TATGGATTGA CCTGACGTGC | 180 |
| CTCAGATGGA TTCTTCCTCC GAAGCCCCCTG GTCATTTCGG AGAATCTGTA ATCTTATTCC | 240 |
| CTTCTTTGGC GAAAATCTGT CAGCTTGAT GTACTCATCC ATCTTCTGAA GCAGCTTCTC | 300 |
| CAGAGTTTGT GGAGGCTTCC TGGCGAAATA TTGGGCTGTA GGTCTTGAC GAAGACCCTT | 360 |
| GATCATGGCC TCAATGACAA TCTCATTGGG CACCGTAGGC GCTTGTGCCC TCAATCGCAA | 420 |
| GAACCTTCGT ACATATGCCT GAAGGTATTC TTCGTGATCT TGTGTGCATT GGAACAGAGC | 480 |
| CTGAGCTGTG ACCGACTTCG TTTGAAAGCC TTGGAAGCTA GTAACCAACA TGTGCTTAAG | 540 |
| CTTCTGCCAC GACGTGATAG TCCCTGGCCG AAGAGAAGAA TACCATGTTT GGGCTACATT | 600 |
| CCGGACTGCC ATGACGAAGG ACTTCGCCAT GACTACAGTG TTGACCCCAT ACGAAGATAT | 660 |

| | | | | | | |
|------------|------------|------------|------------|-------------|------------|------|
| AGTTGCTTCG | TAGCTCATCA | GAAACTGCTT | TGGATCTGAG | TGCCCATCAT | ACATGGGGAG | 720 |
| CTGAGGTGGC | TTGTATGATG | GGGGCCATGG | GGTAGCCTGC | AGTTCTGCTG | CCAAGGGAGA | 780 |
| AGCATCATCA | AAAGTAAAGG | CATCATGATT | AAAATCATCA | TACCATCCAT | CCTCGTTGAA | 840 |
| TAAGCCTTCT | TGACGAAGCT | CCCTGTGTTG | GGGCCTTCGA | TCTTGTTTCAT | CTTGAACAAG | 900 |
| ATGACGCACT | TCTTCAGTGG | CTTCGTCGAT | CTTCTTTTGG | AGATCAGCCA | GTCGCACCAT | 960 |
| CTTCTCCTTC | TTTCTTTGTA | CTTGTTGATG | GATGATCTCC | ATGTCCCTGA | TCTCTTGGTC | 1020 |
| CAACTCCTCC | TCTTGGAGTG | TCAGACTGGT | GGCTTTCCTC | TTCTGGCTTC | GAGCCTCTCG | 1080 |
| AAGAGAAAGA | GTTTCTTGAT | TTGGGTCCAG | CGGCTGCAGT | GCAGTGGTCC | CTGGTGCTGA | 1140 |
| AGCTTTCTTC | GGTGGCATGA | CAAAGGTCAG | TGCTTGCCGA | AGGTGGTCGA | AAAGGGTTCA | 1200 |
| CTAGAGGTGG | GAGCCAATGT | TGGGGACTTC | TCAAGTGCTA | TGAGTTAAGA | ACAAGGCAAC | 1260 |
| ACAAAATGTT | AAATATTAAT | AGCTTTCATC | TTTCGAAGCA | TTATTTCCCT | TTGGGTATAA | 1320 |
| TGATCTTCAG | ACGAAAGAGT | CCTTCATCAT | TGCGATATAT | GTAAATAGAA | GGAGGAGCAT | 1380 |
| ATGAAATGTA | AGAGACAACA | TGAACAATCG | TGTAGCATTG | TTAATTCATC | ATCATTTTAT | 1440 |
| TATTATGGAA | AAATAGAAAC | AATATTGAAT | TACAAATGTA | CCTTTGGCTT | GACAGAAGAT | 1500 |
| AAAAGTACAA | GCTTGACGCA | CGAGCAAGTA | CAAGTCAGTG | TGAACAGTAC | GGGGGTACTG | 1560 |
| TTCATCTATT | TATAGGCACA | GGACACAGCC | TGTGAGAAAT | TACAGTCATG | CCCTTTACAT | 1620 |
| TTACTATTGA | CTTATAGAAA | AATCTATGAG | GACTGGATAG | CCTTTTCCCC | TTTAAGTCGG | 1680 |
| TGCCTTTTTT | CGCGATTAAG | CCGAATCTCC | CTTGCGCATA | GCTTCGGAGC | ATCGGCAACC | 1740 |
| TTCGTCACGA | TCATGCCCTT | CTCATGTGT | ATGCTTTTAA | TCCTGAATTC | GAAGGTACCT | 1800 |
| GTCCATAAAC | CATACTTGGA | AGACATTGTT | AAATTATGTT | TTTGAGGACC | TTCGGAGGAC | 1860 |
| GAAGGCCCCC | AACAGTCGTG | TTTTTGAGGA | CCTTCGGAAG | ATGAAGGCCC | CCAACAAGAC | 1920 |
| CTATCCATAA | AACCAACCTA | TCCACAAAAC | CGACCCCAT | CACCCTTCAT | TTGCCTCACC | 1980 |
| AACAACCCTA | ATTAGGTTGT | TGGTTTAAAT | TTTTTAGGGT | CAATTTGGTC | ATCACCATCC | 2040 |
| ACTGTCACTC | CACAACTCA | ATATCAATAA | ACAGACTCAA | TCACCCAAAC | TGACCATACC | 2100 |
| CATAAAACCG | CCCCACCCTT | CTAGCGCCTC | GCCAGAAACC | AGAAACCCTG | ATTCAGAGTT | 2160 |
| CAAACTTAAA | ACGACCATAA | CTTTCACCTT | GGAAGTCGAA | TCAGGTCCAT | TTTTTTCCAA | 2220 |
| ATCACACAAA | ATTAAATTTC | GCATCCGATA | ATCAAGCCAT | CTCTTCACTA | TGGTTTTAAG | 2280 |
| TGTTGCTCAC | ACTAGTGTAT | TTATGGACTA | ATCACCTGTG | TATCTCATAC | AATAACATAT | 2340 |
| CAGTACATCT | AAGTTGTTAC | TCAATTACCA | AAACCGAATT | ATAGCCTTCG | AAAAAGGTAA | 2400 |
| TCGACTAGTC | ACTCAATTAC | CAAACTAAA | CTTTAGACTT | TCATGTATGA | CATCCAACAT | 2460 |
| GACACTGTAC | TGGACTAAAC | CACCTTTCAA | GCTACACAAG | GAGCAAAAAT | AACTAATTTT | 2520 |

| | |
|---|------|
| CGTAGTTGTA GGAGCTAAAG TATATGTCCA CAACAATAGT TAAGGGAAGC CCCCAGGAC | 2580 |
| TTAAAAGTCC TTTTACCTCT TGAAACTTTT GTCGTGGTCT ACTTTTTCAC TTTAAACTTC | 2640 |
| AAAATTTGAC ATTTTATCAC CCCTTAACTC TTAAAACCAT TTAAATTACA TTCTTACTAG | 2700 |
| ATTATAGATG ATTTTGTGTG TAAAAGTTTT TAAGACATGT TTACACATTG ATTAAAATCA | 2760 |
| TTTGTTCAAT TTCCTAGAGT TAAATCTAAT CTTATTAAAA CTATTAGAGA TACTTTCACG | 2820 |
| AGCTCTAAAT ATTTTTATTT TTTCATTATG GAATTTTGTT AGAATTCCTA TAGACCTTTT | 2880 |
| TTTGTGGTTT AAAAGCCTTG CCATGTTTTT AACAAGTTTT TTTTCTATTT TTTGAAATTT | 2940 |
| TCTTGGAAC CACTTCTAAC CCGGTAGAAG ATTTATTTTG CTACACTTAT ATCTACAACA | 3000 |
| AAATCAACTT ATGAAATTGT CTTGGAAACT ACCTCTAACC CGGTAGAATG AATTGGAATG | 3060 |
| AAAATTAAAC CAACTTACGG AATCGCCCAA CATATGTCGA TTAAAGTGGA TATGGATACA | 3120 |
| TATGAAGAAG CCCTAGAGAT AATCTAAATG GTTTCAGAAT TGAGGGTTAT TTTTGAAGT | 3180 |
| TTGATGGGAA GATAAGACCA TAACGGTAGT TCACAGAGAT AAAAGGGTTA TTTTTTTCAG | 3240 |
| AAATATTTGT GCTGCAATTG ATCCTGTGCC TCAAATTCAG CCTGCAACCA AGGCCAGGTT | 3300 |
| CTAGAGCGAA CAAGGCCAC GTCACCCGTG GCCCGTCAGG CGAAGCAGGT CTTGTGCAGA | 3360 |
| CTTTGAGAGG GATTGGATAT CAACGGAACC AATCACGCAC GGCAATGCGA TTCCCAGCCC | 3420 |
| ACCTGTAACG TTCCAGTGGG CCATCCTTAA CTCCAAGCCC AACGGCCCTA CCCCATCTCG | 3480 |
| TCGTGTCATC CACTCCGCCG CACAGGCGCT CAGCTCCGCA ACGCCGCCGG AAATGGTCGC | 3540 |
| CGCCACAGCC ACCGCCATGG CCACCGCTGC ATCGCCGCTA CTCAACGGGA CCCGAATACC | 3600 |
| TGCGCGGCTC CGCCATCGAG GACTCAGCGT GCGCTGCGCT GCTGTGGCGG GCGGCGCGGC | 3660 |
| CGAGGCACCG GCATCCACCG GCGCGCGGCT GTCCGCGGAC TGCGTTGTGG TGGGCGGAGG | 3720 |
| CATCAGTGGC CTCTGCACCG CGCAGGCGCT GGCCACGCGG CACGGCGTCG GGGACGTGCT | 3780 |
| TGTCACGGAG GCCCGCGCCC GCCCCGGCGG CAACATTACC ACCGTCGAGC GCCCCGAGGA | 3840 |
| AGGGTACC | 3848 |

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1826 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Gossypium hirsutum* (cotton)

(vii) IMMEDIATE SOURCE:
 (B) CLONE: pWDC-15 (NRRL B-21594)

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: 31..1647
 (D) OTHER INFORMATION: /product= "Cotton protox-1 coding sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

| | |
|---|------|
| CCTCTCGCTC GCCTGGCCCC ACCACCAATC ATGACGGCTC TAATCGACCT TTCTCTTCTC | 60 |
| CGTTCCTCGC CCTCCGTTTC CCCTTTCTCC ATACCCACC ACCAGCATCC GCCCCGCTTT | 120 |
| CGTAAACCTT TCAAGCTCCG ATGCTCCCTC GCCGAGGGTC CCACGATTTC CTCATCTAAA | 180 |
| ATCGACGGGG GAGAATCATC CATCGCGGAT TGCCTCATCG TTGGAGGTGG TATCAGTGGA | 240 |
| CTTTGCATTG CTCAAGCTCT CGCCACCAAG CACCGTGACG TCGCTTCCAA TGTGATTGTG | 300 |
| ACGGAGGCCA GAGACCGTGT TGGTGGCAAC ATCACTACCG TTGAGAGAGA TGGATATCTG | 360 |
| TGGAAGAAG GCCCAACAG TTTTCAGCCC TCCGATCCTA TTCTAACCAT GGCCGTGGAT | 420 |
| AGTGGATTGA AGGACGATTT GGTTTTAGGT GACCTAATG CACCGCGATT TGTAATATGG | 480 |
| GAGGAAAAC TAAGGCTGT GCCCTCCAAG CCAACCGACT TGCCGTTTTT TGATTTGATG | 540 |
| AGCATTGCTG GAAACTTAG GGCTGGGTTT GGGGCTATTG GCATTCGGCC TCCCCCTCCG | 600 |
| GGTTATGAAG AATCGGTGGA GGAGTTTGTG CGCCGTAATC TTGGTGCTGA GGTTTTTGAA | 660 |
| CGCTTTATTG AACCATTTTG TTCAGGTGTT TATGCAGGGG ATCCTTCAAA ATTAAGCATG | 720 |
| AAAGCAGCAT TTGGAAGAGT ATGGAAGCTA GAAGAGATTG GTGGCAGCAT CATTGGTGGC | 780 |
| ACTTTCAAGA CAATCCAGGA GAGAAATAAG ACACCTAAGC CACCCAGAGA CCCGCGTCTG | 840 |
| CCAAAACCGA AGGGCCAAAC AGTTGGATCT TTTAGGAAGG GACTTACCAT GCTGCCTGAG | 900 |
| GCAATTGCTA ACAGTTTGGG TAGCAATGTA AAATTATCTT GGAAGCTTTC CAGTATTACC | 960 |
| AAATTGGGCA ATGGAGGGTA TAACTTGACA TTTGAAACAC CTGAAGGAAT GGTATCTCTT | 1020 |
| CAGAGTAGAA GTGTTGTAAT GACCATTCCA TCCCATGTTG CCAGTAACTT GTTGCATCCT | 1080 |
| CTCTCGGCTG CTGCTGCAGA TGCATTATCC CAATTTTATT ATCCTCCAGT TGCATCAGTC | 1140 |
| ACAGTCTCCT ATCCAAAAGA AGCCATTCTG AAAGAATGTT TGATTGATGG TGAACCTAAG | 1200 |
| GGGTTTGGCC AGTTGCACCC ACGCAGCCAA GGAATTGAAA CTTTAGGGAC GATATACAGT | 1260 |
| TCATCACTTT TCCCAATCG AGCTCCATCT GGCAGGGTGT TGCTCTTGAA CTACATAGGA | 1320 |
| GGAGCTACCA AACTGGAAT TTTGTCCAAG ACTGAAGGGG AACTTGTAGA AGCAGTTGAT | 1380 |

```

CGTGATTTGA GAAAAATGCT TATAAATCCT AATGCAAAGG ATCCTCTTGT TTTGGGTGTA      1440
AGAGTATGGC CAAAAGCCAT TCCACAGTTC TTGGTTGGTC ATTTGGATCT CCTTGATAGT      1500
GCAAAAATGG CTCTCAGGGA TTCTGGGTTT CATGGACTGT TTCTTGGGGG CAACTATGTA      1560
TCTGGTGTGG CATTAGGACG GTGTGTGGAA GGTGCTTACG AGGTTGCAGC TGAAGTGAAG      1620
GAATTCCTGT CACAATATGC ATACAAATAA TATTGAAATT CTTGTCAGGC TGCAAATGTA      1680
GAAGTCAGTT ATTGGATAGT ATCTCTTTAG CTAAAAAATT GGGTAGGGTT TTTTTTGTTA      1740
GTTCTTGAC CACTTTTTGG GGTTCATT AGAACTTCAT ATTTGTATAT CATGTTGCAA      1800
TATCAAAAAA AAAAAAAAAA AAAAAA      1826

```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 539 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```

Met Thr Ala Leu Ile Asp Leu Ser Leu Leu Arg Ser Ser Pro Ser Val
1          5          10          15

Ser Pro Phe Ser Ile Pro His His Gln His Pro Pro Arg Phe Arg Lys
20          25          30

Pro Phe Lys Leu Arg Cys Ser Leu Ala Glu Gly Pro Thr Ile Ser Ser
35          40          45

Ser Lys Ile Asp Gly Gly Glu Ser Ser Ile Ala Asp Cys Val Ile Val
50          55          60

Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala Thr Lys
65          70          75          80

His Arg Asp Val Ala Ser Asn Val Ile Val Thr Glu Ala Arg Asp Arg
85          90          95

Val Gly Gly Asn Ile Thr Thr Val Glu Arg Asp Gly Tyr Leu Trp Glu
100         105         110

Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Ile Leu Thr Met Ala
115         120         125

Val Asp Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Asn Ala
130         135         140

Pro Arg Phe Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys
145         150         155         160

```

Pro Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Ala Gly Lys Leu
 165 170 175
 Arg Ala Gly Phe Gly Ala Ile Gly Ile Arg Pro Pro Pro Gly Tyr
 180 185 190
 Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val
 195 200 205
 Phe Glu Arg Phe Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp
 210 215 220
 Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Arg Val Trp Lys Leu
 225 230 235 240
 Glu Glu Ile Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys Thr Ile Gln
 245 250 255
 Glu Arg Asn Lys Thr Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Lys
 260 265 270
 Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Thr Met Leu
 275 280 285
 Pro Glu Ala Ile Ala Asn Ser Leu Gly Ser Asn Val Lys Leu Ser Trp
 290 295 300
 Lys Leu Ser Ser Ile Thr Lys Leu Gly Asn Gly Gly Tyr Asn Leu Thr
 305 310 315 320
 Phe Glu Thr Pro Glu Gly Met Val Ser Leu Gln Ser Arg Ser Val Val
 325 330 335
 Met Thr Ile Pro Ser His Val Ala Ser Asn Leu Leu His Pro Leu Ser
 340 345 350
 Ala Ala Ala Ala Asp Ala Leu Ser Gln Phe Tyr Tyr Pro Pro Val Ala
 355 360 365
 Ser Val Thr Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu
 370 375 380
 Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Ser Gln
 385 390 395 400
 Gly Ile Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn
 405 410 415
 Arg Ala Pro Ser Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala
 420 425 430
 Thr Asn Thr Gly Ile Leu Ser Lys Thr Glu Gly Glu Leu Val Glu Ala
 435 440 445
 Val Asp Arg Asp Leu Arg Lys Met Leu Ile Asn Pro Asn Ala Lys Asp
 450 455 460
 Pro Leu Val Leu Gly Val Arg Val Trp Pro Lys Ala Ile Pro Gln Phe
 465 470 475 480

Leu Val Gly His Leu Asp Leu Leu Asp Ser Ala Lys Met Ala Leu Arg
 485 490 495
 Asp Ser Gly Phe His Gly Leu Phe Leu Gly Gly Asn Tyr Val Ser Gly
 500 505 510
 Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Val Ala Ala Glu
 515 520 525
 Val Lys Glu Phe Leu Ser Gln Tyr Ala Tyr Lys
 530 535

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1910 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Beta vulgaris (Sugar Beet)
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: pWDC-16 (NRRL B-21595N)
- (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION: 1..1680
 - (D) OTHER INFORMATION: /product= "Sugar Beet protox-1 coding sequence"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

| | |
|--|-----|
| ATGAAATCAA TGGCGTTATC AAAGTGCATT CCACAGACAC AGTGCATGCC ATTGCGCAGC | 60 |
| AGCGGGCATT ACAGGGGTAA TTGTATCATG TTGTCAATTC CATGTAGTTT AATTGGAAGA | 120 |
| CGAGGTTATT ATTCACATAA GAAGAGGAGG ATGAGCATGA GTTGCAGCAC AAGCTCAGGC | 180 |
| TCAAAGTCAG CGGTTAAAGA AGCAGGATCA GGATCAGGTG CAGGAGGATT GCTAGACTGC | 240 |
| GTAATCGTTG GAGGTGGAAT TAGCGGGCTT TGCATCGCGC AGGCTCTTTG TACAAAACAC | 300 |
| TCCTCTTCCT CTTTATCCCC AAATTTTATA GTTACAGAGG CCAAAGACAG AGTTGGCGGC | 360 |
| AACATCGTCA CTGTGGAGGC CGATGGCTAT ATCTGGGAGG AGGGACCCAA TAGCTTCCAG | 420 |
| CCTTCCGACG CGGTGCTCAC CATGGCGGTC GACAGTGGCT TGAAAGATGA GTTGGTGCTC | 480 |
| GGAGATCCCA ATGCTCCTCG CTTTGTGCTA TGGAAATGACA AATTAAGGCC CGTACCTTCC | 540 |
| AGTCTCACCG ACCTCCCTTT CTTGACCTC ATGACCATTC CGGGCAAGAT TAGGGCTGCT | 600 |

```

CTTGGTGCTC TCGGATTTTCG CCCTTCTCCT CCACCTCATG AGGAATCTGT TGAACACTTT      660
GTGCGTCGTA ATCTCGGAGA TGAGGTCTTT GAACGCTTGA TTGAACCCCTT TTGTTCAGGT      720
GTGTATGCCG GTGATCCTGC CAAGCTGAGT ATGAAAGCTG CTTTGGGAA GGTCTGGAAG      780
TTGGAGCAAA AGGGTGGCAG CATAATTGGT GGCACCTCTCA AAGCTATACA GGAAAGAGGG      840
AGTAATCCTA AGCCGCCCCG TGACCAGCGC CTCCTTAAAC CAAAGGGTCA GACTGTTGGA      900
TCCTTTAGAA AGGGACTCGT TATGTTGCCT ACCGCCATTT CTGCTCGACT TGGCAGTAGA      960
GTGAAACTAT CTTGGACCCT TTCTAGTATC GTAAAGTCAC TCAATGGAGA ATATAGTCTG     1020
ACTTATGATA CCCAGATGG CTTGGTTTCT GTAAGAACCA AAAGTGTGTG GATGACTGTT     1080
CCATCATATG TTGCAAGTAG GCTTCTTCGT CCACTTTCAG ACTCTGCTGC AGATTCTCTT     1140
TCAAAATTTT ACTATCCACC AGTTGCAGCA GTGTCACTTT CCTATCCTAA AGAAGCGATC     1200
AGATCAGAAT GCTTGATTAA TGGTGAACCT CAAGGTTTCG GGCAACTACA TCCCCGCAGT     1260
CAGGGTGTGG AAACCTTGGG AACAAATTTAT AGTTCGTCTC TTTTCCCTGG TCGAGCACCA     1320
CCTGGTAGGA TCTTGATCTT GAGCTACATC GGAGGTGCTA AAAATCCTGG CATATTAAAC     1380
AAGTCGAAAG ATGAACCTGC CAAGACAGTT GACAAGGACC TGAGAAGAAT GCTTATAAAT     1440
CCTGATGCAA AACTTCCTCG TGTACTGGGT GTGAGAGTAT GGCCTCAAGC AATACCCCAG     1500
TTTTCTATTG GGCACCTTGA TCTGCTCGAT GCTGCAAAAAG CTGCTCTGAC AGATACAGGG     1560
GTCAAAGGAC TGTTTCTTGG TGGCAACTAT GTTTCAGGTG TTGCCTTGGG GCGGTGTATA     1620
GAGGGTGCTT ATGAGTCTGC AGCTGAGGTA GTAGATTTC TCTCACAGTA CTCAGACAAA     1680
TAGAGCTTCA GCATCCTGTG TAATTCAACA CAGGCCTTTT TGTATCTGTT GTGCGCGCAT     1740
GTAGTCTGGT CGTGGTGCTA GGATTGATTA GTTGCTCTGC TGTGTGATCC ACAAGAATTT     1800
TGATGGAATT TTTCCAGATG TGGGCATTAT ATGTTGCTGT CTTATAAATC CTTAATTTGT     1860
ACGTTTAGTG AATTACACCG CATTTGATGA CTAACAAAAA AAAAAA      1910

```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 560 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

Met Lys Ser Met Ala Leu Ser Asn Cys Ile Pro Gln Thr Gln Cys Met
1           5           10           15

```

Pro Leu Arg Ser Ser Gly His Tyr Arg Gly Asn Cys Ile Met Leu Ser
 20 25 30
 Ile Pro Cys Ser Leu Ile Gly Arg Arg Gly Tyr Tyr Ser His Lys Lys
 35 40 45
 Arg Arg Met Ser Met Ser Cys Ser Thr Ser Ser Gly Ser Lys Ser Ala
 50 55 60
 Val Lys Glu Ala Gly Ser Gly Ser Gly Ala Gly Gly Leu Leu Asp Cys
 65 70 75 80
 Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu
 85 90 95
 Cys Thr Lys His Ser Ser Ser Ser Leu Ser Pro Asn Phe Ile Val Thr
 100 105 110
 Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Val Thr Val Glu Ala Asp
 115 120 125
 Gly Tyr Ile Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Ala
 130 135 140
 Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu
 145 150 155 160
 Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Asn Asp Lys Leu Arg
 165 170 175
 Pro Val Pro Ser Ser Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Thr
 180 185 190
 Ile Pro Gly Lys Ile Arg Ala Ala Leu Gly Ala Leu Gly Phe Arg Pro
 195 200 205
 Ser Pro Pro Pro His Glu Glu Ser Val Glu His Phe Val Arg Arg Asn
 210 215 220
 Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly
 225 230 235 240
 Val Tyr Ala Gly Asp Pro Ala Lys Leu Ser Met Lys Ala Ala Phe Gly
 245 250 255
 Lys Val Trp Lys Leu Glu Gln Lys Gly Gly Ser Ile Ile Gly Gly Thr
 260 265 270
 Leu Lys Ala Ile Gln Glu Arg Gly Ser Asn Pro Lys Pro Pro Arg Asp
 275 280 285
 Gln Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys
 290 295 300
 Gly Leu Val Met Leu Pro Thr Ala Ile Ser Ala Arg Leu Gly Ser Arg
 305 310 315 320
 Val Lys Leu Ser Trp Thr Leu Ser Ser Ile Val Lys Ser Leu Asn Gly
 325 330 335

Glu Tyr Ser Leu Thr Tyr Asp Thr Pro Asp Gly Leu Val Ser Val Arg
 340 345 350
 Thr Lys Ser Val Val Met Thr Val Pro Ser Tyr Val Ala Ser Arg Leu
 355 360 365
 Leu Arg Pro Leu Ser Asp Ser Ala Ala Asp Ser Leu Ser Lys Phe Tyr
 370 375 380
 Tyr Pro Pro Val Ala Ala Val Ser Leu Ser Tyr Pro Lys Glu Ala Ile
 385 390 395 400
 Arg Ser Glu Cys Leu Ile Asn Gly Glu Leu Gln Gly Phe Gly Gln Leu
 405 410 415
 His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser
 420 425 430
 Ser Leu Phe Pro Gly Arg Ala Pro Pro Gly Arg Ile Leu Ile Leu Ser
 435 440 445
 Tyr Ile Gly Gly Ala Lys Asn Pro Gly Ile Leu Asn Lys Ser Lys Asp
 450 455 460
 Glu Leu Ala Lys Thr Val Asp Lys Asp Leu Arg Arg Met Leu Ile Asn
 465 470 475 480
 Pro Asp Ala Lys Leu Pro Arg Val Leu Gly Val Arg Val Trp Pro Gln
 485 490 495
 Ala Ile Pro Gln Phe Ser Ile Gly His Phe Asp Leu Leu Asp Ala Ala
 500 505 510
 Lys Ala Ala Leu Thr Asp Thr Gly Val Lys Gly Leu Phe Leu Gly Gly
 515 520 525
 Asn Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala Tyr
 530 535 540
 Glu Ser Ala Ala Glu Val Val Asp Phe Leu Ser Gln Tyr Ser Asp Lys
 545 550 555 560

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1784 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus (oilseed rape)

(vii) IMMEDIATE SOURCE:

(B) CLONE: pWDC-17 (NRRL B-21615)

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION: 47..1654

(D) OTHER INFORMATION: /product= "Oilseed rape protox-1 coding sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

| | |
|--|------|
| GGGCCCCCCC CAAAATTGAG GATTCTCCTT CTCGCGGGCG ATCGCCATGG ATTTATCTCT | 60 |
| TCTCCGTCCG CAGCCATTCC TATCGCCATT CTCAAATCCA TTTCTCTCGGT CGCGTCCCTA | 120 |
| CAAGCCTCTC AACCTCCGTT GCTCCGTATC CGGTGGATCC GTCGTCGGCT CTTCTACAAT | 180 |
| CGAAGGCGGA GGAGGAGGTA AAACCGTCAC GCGGACTGC GTGATCGTCG GCGGAGGAAT | 240 |
| CAGCGGCCTG TGCATTGCGC AAGCGCTCGT GACGAAGCAC CCAGACGCTG CAAAGAATGT | 300 |
| GATGGTGACG GAGGCGAAGG ACCGTGTGGG AGGGAATATC ATCACGCGAG AGGAGCAAGG | 360 |
| GTTTCTATGG GAAGAAGGTC CCAATAGCTT TCAGCCGTCT GATCCTATGC TCACTATGGT | 420 |
| GGTAGATAGT GGTTTGAAAG ATGATCTAGT CTTGGGAGAT CCTACTGCTC CGAGGTTTGT | 480 |
| GTTGTGGAAT GGGAAGCTGA GGCCGGTTCC GTCGAAGCTA ACTGACTTGC CTTTCTTTGA | 540 |
| CTTGATGAGT ATTGGAGGGA AGATTAGAGC TGGGTTTGGT GCCATTGGTA TTCGACCTTC | 600 |
| ACCTCCGGGT CGTGAGGAAT CAGTGGAAGA GTTTGTAAGG CGTAATCTTG GTGATGAGGT | 660 |
| TTTTGAGCGC TTGATTGAAC CCTTTTGCTC AGGTGTTTAT GCGGGAGATC CTGCGAAACT | 720 |
| GAGTATGAAA GCAGCTTTTG GGAAGGTTTG GAAGCTAGAG GAGAATGGTG GGAGCATCAT | 780 |
| TGGTGGTGCT TTTAAGGCAA TTCAAGCGAA AAATAAAGCT CCCAAGACAA CCCGAGATCC | 840 |
| GCGTCTGCCA AAGCCAAAGG GCCAACTGT TGGTTCTTTC AGGAAAGGAC TCACAATGCT | 900 |
| GCCAGAGGCA ATCTCCGCAA GGTTGGGTGA CAAGGTGAAA GTTTCTTGGA AGCTCTCAAG | 960 |
| TATCACTAAG CTGGCCAGCG GAGAATATAG CTTAACTTAC GAAACTCCGG AGGGTATAGT | 1020 |
| CACTGTACAG AGCAAAAGTG TAGTGATGAC TGTGCCATCT CATGTTGCTA GTAGTCTCTT | 1080 |
| GCGCCCTCTC TCTGATTCTG CAGCTGAAGC GCTCTCAAAA CTCTACTATC CGCCAGTTGC | 1140 |
| AGCCGTATCC ATCTCATACG CGAAAGAAGC AATCCGAAGC GAATGCTTAA TAGATGGTGA | 1200 |
| ACTAAAAGGG TTCGGCCAGT TGCATCCACG CACGCAAAAA GTGGAAACTC TTGGAACAAT | 1260 |
| ATACAGTTCA TCGCTCTTTC CCAACCGAGC ACCGCCTGGA AGAGTATTGC TATTGAACTA | 1320 |
| CATCGGTGGA GCTACCAACA CTGGGATCTT ATCAAAGTCG GAAGGTGAGT TAGTGGAAGC | 1380 |
| AGTAGATAGA GACTTGAGGA AGATGCTGAT AAAGCCAAGC TCGACCGATC CACTTGTA | 1440 |
| TGGAGTAAAA TTATGGCCTC AAGCCATTCC TCAGTTTCTG ATAGGTCACA TTGATTTGGT | 1500 |


```

AGACGCAGCG AAAGCATCGC TCTCGTCATC TGGTCATGAG GGCTTATTCT TGGGTGGAAA      1560
TTACGTTGCC GGTGTAGCAT TGGGTCGGTG TGTGGAAGGT GCTTATGAAA CTGCAACCCA      1620
AGTGAATGAT TTCATGTCAA GGTATGCTTA CAAGTAATGT AACGCAGCAA CGATTTGATA      1680
CTAAGTAGTA GATTTTGCAG TTTTGACTTT AAGAACACTC TGTTTGTGAA AAATTCAAGT      1740
CTGTGATTGA GTAAATTTAT GTATTATTAC TAAAAA AAAA                          1784

```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 536 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Met Asp Leu Ser Leu Leu Arg Pro Gln Pro Phe Leu Ser Pro Phe Ser
1           5           10           15
Asn Pro Phe Pro Arg Ser Arg Pro Tyr Lys Pro Leu Asn Leu Arg Cys
          20           25           30
Ser Val Ser Gly Gly Ser Val Val Gly Ser Ser Thr Ile Glu Gly Gly
          35           40           45
Gly Gly Gly Lys Thr Val Thr Ala Asp Cys Val Ile Val Gly Gly Gly
          50           55           60
Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Val Thr Lys His Pro Asp
65           70           75           80
Ala Ala Lys Asn Val Met Val Thr Glu Ala Lys Asp Arg Val Gly Gly
          85           90           95
Asn Ile Ile Thr Arg Glu Glu Gln Gly Phe Leu Trp Glu Glu Gly Pro
          100          105          110
Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr Met Val Val Asp Ser
          115          120          125
Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Thr Ala Pro Arg Phe
          130          135          140
Val Leu Trp Asn Gly Lys Leu Arg Pro Val Pro Ser Lys Leu Thr Asp
          145          150          155          160
Leu Pro Phe Phe Asp Leu Met Ser Ile Gly Gly Lys Ile Arg Ala Gly
          165          170          175
Phe Gly Ala Ile Gly Ile Arg Pro Ser Pro Pro Gly Arg Glu Glu Ser
          180          185          190

```

Val Glu Glu Phe Val Arg Arg Asn Leu Gly Asp Glu Val Phe Glu Arg
 195 200 205
 Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ala Lys
 210 215 220
 Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Lys Leu Glu Glu Asn
 225 230 235 240
 Gly Gly Ser Ile Ile Gly Gly Ala Phe Lys Ala Ile Gln Ala Lys Asn
 245 250 255
 Lys Ala Pro Lys Thr Thr Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly
 260 265 270
 Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Thr Met Leu Pro Glu Ala
 275 280 285
 Ile Ser Ala Arg Leu Gly Asp Lys Val Lys Val Ser Trp Lys Leu Ser
 290 295 300
 Ser Ile Thr Lys Leu Ala Ser Gly Glu Tyr Ser Leu Thr Tyr Glu Thr
 305 310 315 320
 Pro Glu Gly Ile Val Thr Val Gln Ser Lys Ser Val Val Met Thr Val
 325 330 335
 Pro Ser His Val Ala Ser Ser Leu Leu Arg Pro Leu Ser Asp Ser Ala
 340 345 350
 Ala Glu Ala Leu Ser Lys Leu Tyr Tyr Pro Pro Val Ala Ala Val Ser
 355 360 365
 Ile Ser Tyr Ala Lys Glu Ala Ile Arg Ser Glu Cys Leu Ile Asp Gly
 370 375 380
 Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Thr Gln Lys Val Glu
 385 390 395 400
 Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
 405 410 415
 Pro Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr
 420 425 430
 Gly Ile Leu Ser Lys Ser Glu Gly Glu Leu Val Glu Ala Val Asp Arg
 435 440 445
 Asp Leu Arg Lys Met Leu Ile Lys Pro Ser Ser Thr Asp Pro Leu Val
 450 455 460
 Leu Gly Val Lys Leu Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly
 465 470 475 480
 His Ile Asp Leu Val Asp Ala Ala Lys Ala Ser Leu Ser Ser Ser Gly
 485 490 495
 His Glu Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
 500 505 510
 Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala Thr Gln Val Asn Asp

515

520

525

Phe Met Ser Arg Tyr Ala Tyr Lys
530 535

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1224 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Oryza sativa (rice)

(vii) IMMEDIATE SOURCE:

(B) CLONE: pWDC-18 (NRRL B-21648)

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..936
- (D) OTHER INFORMATION: /product= "Rice protox-1 partial coding sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

| | |
|---|-----|
| CGGGCTTTGA AGGCTGCATT TGGGAAGGTG TGGAGGCTGG AGGATACTGG AGGTAGCATT | 60 |
| ATTGGTGGAA CCATCAAGAC AATCCAGGAG AGGGGGAAAA ACCCCAAACC GCCGAGGGAT | 120 |
| CCCCGCCTTC CAACGCCAAA GGGGCAGACA GTTGCATCTT TCAGGAAGGG TCTGACTATG | 180 |
| CTCCCGGATG CTATTACATC TAGGTTGGGT AGCAAAGTCA AACTTTCATG GAAGTTGACA | 240 |
| AGCATTACAA AGTCAGACAA CAAAGGATAT GCATTAGTGT ATGAAACACC AGAAGGGGTG | 300 |
| GTCTCGGTGC AAGCTAAAAC TGTGTGCATG ACCATCCCAT CATATGTTGC TAGTGATATC | 360 |
| TTGCGGCCAC TTTCAAGTGA TGCAGCAGAT GCTCTGTCAA TATTCTATTA TCCACCAGTT | 420 |
| GCTGCTGTAA CTGTTTCATA TCCAAAAGAA GCAATTAGAA AAGAATGCTT AATTGACGGA | 480 |
| GAGCTCCAGG GTTTCGGCCA GCTGCATCCG CGTAGTCAGG GAGTTGAGAC TTTAGGAACA | 540 |
| ATATATAGCT CATCACTCTT TCCAAATCGT GCTCCAGCTG GAAGGGTGTT ACTTCTGAAC | 600 |
| TACATAGGAG GTTCTACAAA TACAGGGATT GTTTCCAAGA CTGAAAGTGA GCTGGTAGAA | 660 |
| GCAGTTGACC GTGACCTCAG GAAGATGCTG ATAAATCCTA GAGCAGTGA CCCTTTGGTC | 720 |
| CTTGCGTCC GGGTATGGCC ACAAGCCATA CCACAGTTCC TCATTGGCCA TCTTGATCAT | 780 |
| CTTGAGGCTG CAAAATCTGC CCTGGGCAAA GGTGGGTATG ATGGATTGTT CCTCGGAGGG | 840 |

```

AACTATGTTG CAGGAGTTGC CCTGGGCCGA TGC GTTGAAG GTGCATATGA GAGTGCCTCA      900
CAAATATCTG ACTACTTGAC CAAGTACGCC TACAAGTGAT CAAAGTTGGC CTGCTCCTTT      960
TGGCACATAG ATGTGAGGCT TCTAGCAGCA AAAATTTTCAT GGGCATCTTT TTATCCTGAT     1020
TCTAATTAGT TAGAATTTAG AATTGTAGAG GAATGTTCCA TTTGCAGTTC ATAATAGTTG     1080
TTCAGATTTT AGCCATTCAA TTTGTGCAGC CATTACTAT ATGTAGTATG ATCTTGTAAG     1140
TACTACTAAG AACAAATCAA TTATATTTTC CTGCAAGTGA CATCTTAATC GTCAGCAAAT     1200
CCAGTTACTA GTAAAAAAAA AAAA                                           1224

```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 312 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```

Arg Ala Leu Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Asp Thr
1           5           10           15
Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly
20          25          30
Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Thr Pro Lys Gly
35          40          45
Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Thr Met Leu Pro Asp Ala
50          55          60
Ile Thr Ser Arg Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr
65          70          75          80
Ser Ile Thr Lys Ser Asp Asn Lys Gly Tyr Ala Leu Val Tyr Glu Thr
85          90          95
Pro Glu Gly Val Val Ser Val Gln Ala Lys Thr Val Val Met Thr Ile
100         105         110
Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Ser Asp Ala
115        120        125
Ala Asp Ala Leu Ser Ile Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr
130        135        140
Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly
145        150        155        160
Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu

```

(A) LENGTH: 1590 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

TCCACCGTCG AGCGCCCCGA GGAAGGGTAC CTCTGGGAGG AGGGTCCCAA CAGCTTCCAG 60
CCATCCGACC CCGTTCTCTC CATGGCCGTG GACAGCGGGC TGAAGGATGA CCTGGTTTTT 120

| | | | | | | |
|-------------|-------------|------------|------------|------------|------------|------|
| GGGGACCCCA | ACGCGCCACG | GTTCGTGCTG | TGGGAGGGGA | AGCTGAGGCC | CGTGCCATCC | 180 |
| AAGCCC GCCG | ACCTCCCGTT | CTTCGATCTC | ATGAGCATCC | CTGGCAAGCT | CAGGGCCGGT | 240 |
| CTCGGCGCGC | TTGGCATCCG | CCCGCCTGCT | CCAGGCCGCG | AGGAGTCAGT | GGAGGAGTTT | 300 |
| GTGCGCCGCA | ACCTCGGTGC | TGAGGTCTTT | GAGCGCCTAA | TTGAGCCTTT | CTGCTCAGGT | 360 |
| GTCTATGCTG | GCGATCCTTC | CAAGCTCAGT | ATGAAGGCTG | CATTTGGGAA | GGTGTGGCGG | 420 |
| TTAGAAGAAG | CTGGAGGTAG | TATTATTGGT | GGAACCATCA | AGACGATTCA | GGAGAGGGGC | 480 |
| AAGAATCCAA | AACCACCGAG | GGATCCCCGC | CTTCCGAAGC | CAAAAGGGCA | GACAGTTGCA | 540 |
| TCTTTCAGGA | AGGGTCTTGC | CATGCTTCCA | AATGCCATCA | CATCCAGCTT | GGGTAGTAAA | 600 |
| GTCAAACTAT | CATGGAAACT | CACGAGCATG | ACAAAATCAG | ATGGCAAGGG | GTATGTTTTG | 660 |
| GAGTATGAAA | CACCAGAAGG | GGTTGTTTTG | GTGCAGGCTA | AAAGTGTTAT | CATGACCATT | 720 |
| CCATCATATG | TTGCTAGCGA | CATTTTGCCT | CCACTTTCAG | GTGATGCTGC | AGATGTTCTA | 780 |
| TCAAGATTCT | ATTATCCACC | AGTTGCTGCT | GTAACGGTTT | CGTATCCAAA | GGAAGCAATT | 840 |
| AGAAAAGAAT | GCTTAATTGA | TGGGGAAGTC | CAGGGTTTTG | GCCAGTTGCA | TCCACGTAGT | 900 |
| CAAGGAGTTG | AGACATTAGG | AACAATATAC | AGCTCATCAC | TCTTTCCAAA | TCGTGCTCCT | 960 |
| GCTGGTAGGG | TGTTACTTCT | AAACTACATA | GGAGGTGCTA | CAAACACAGG | AATTGTTTCC | 1020 |
| AAGACTGAAA | GTGAGCTGGT | AGAAGCAGTT | GACCGTGACC | TCCGAAAAAT | GCTTATAAAT | 1080 |
| CCTACAGCAG | TGGACCCTTT | AGTCCTTGGT | GTCCGAGTTT | GGCCACAAGC | CATACCTCAG | 1140 |
| TTCTTGGTAG | GACATCTTGA | TCTTCTGGAG | GCCGCAAAAT | CTGCCCTGGA | CCAAGGTGGC | 1200 |
| TATAATGGGC | TGTTCCCTAGG | AGGGAAGTAT | GTTGCAGGAG | TTGCCCTGGG | CAGATGCATT | 1260 |
| GAGGGCGCAT | ATGAGAGTGC | CGCGCAAATA | TATGACTTCT | TGACCAAGTA | CGCCTACAAG | 1320 |
| TGATGGAAGA | AGTGGAGCGC | TGCTTGTTAA | TTGTTATGTT | GCATAGATGA | GGTGAGACCA | 1380 |
| GGAGTAGTAA | AAGGCGTCAC | GAGTATTTTT | CATTCTTATT | TTGTAAATTG | CACTTCTGTT | 1440 |
| TTTTTTTCCT | GTCAGTAATT | AGTTAGATTT | TAGTTATGTA | GGAGATTGTT | GTGTTCACTG | 1500 |
| CCCTACAAAA | GAATTTTTAT | TTTGCATTCT | TTTATGAGAG | CTGTGCAGAC | TTATGTAACG | 1560 |
| TTTTACTGTA | AGTATCAACA | AAATCAAATA | | | | 1590 |

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 440 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```

Ser Thr Val Glu Arg Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro
1           5           10           15
Asn Ser Phe Gln Pro Ser Asp Pro Val Leu Ser Met Ala Val Asp Ser
20           25           30
Gly Leu Lys Asp Asp Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe
35           40           45
Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp
50           55           60
Leu Pro Phe Phe Asp Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly
65           70           75           80
Leu Gly Ala Leu Gly Ile Arg Pro Pro Ala Pro Gly Arg Glu Glu Ser
85           90           95
Val Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg
100          105          110
Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys
115          120          125
Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Ala
130          135          140
Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly
145          150          155          160
Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly
165          170          175
Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Ala Met Leu Pro Asn Ala
180          185          190
Ile Thr Ser Ser Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr
195          200          205
Ser Met Thr Lys Ser Asp Gly Lys Gly Tyr Val Leu Glu Tyr Glu Thr
210          215          220
Pro Glu Gly Val Val Leu Val Gln Ala Lys Ser Val Ile Met Thr Ile
225          230          235          240
Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Gly Asp Ala
245          250          255
Ala Asp Val Leu Ser Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr
260          265          270
Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly
275          280          285
Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu
290          295          300

```

```

Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
305                      310                      315                      320

Ala Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr
                      325                      330                      335

Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg
                      340                      345                      350

Asp Leu Arg Lys Met Leu Ile Asn Pro Thr Ala Val Asp Pro Leu Val
                      355                      360                      365

Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly
370                      375                      380

His Leu Asp Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Gln Gly Gly
385                      390                      395                      400

Tyr Asn Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
                      405                      410                      415

Gly Arg Cys Ile Glu Gly Ala Tyr Glu Ser Ala Ala Gln Ile Tyr Asp
                      420                      425                      430

Phe Leu Thr Lys Tyr Ala Tyr Lys
435                      440

```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "maize protox-1 intron sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```

GTACGCTCCT CGCTGGCGCC GCAGCGTCTT CTTCTCAGAC TCATGCGCAG CCATGGAATT      60
GAGATGCTGA ATGGATTTTA TACGCGCGCG CAG                                     93

```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2606 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Beta vulgaris (sugar beet)

(vii) IMMEDIATE SOURCE:
 (B) CLONE: pWDC-20 (NRRL B-21650)

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: 2601..2606
 (D) OTHER INFORMATION: /note= "SalI site"

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: complement (1..538)
 (D) OTHER INFORMATION: /note= "partial cDNA of sugar beet
 protox-1"

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: 539..2606
 (D) OTHER INFORMATION: /note= "sugar beet protox-1
 promoter region (partial sequence of the ~ 3 kb PstI-SalI fragment subcloned
 from pWDC-20)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

| | |
|---|-----|
| CTGCAGGGGG AGGGAAAGAG AGACCGCGAC GGTGAGGGAG GGGAGACCGC GACGGTGAGG | 60 |
| GAGGGGAGAA CGCGACGGTG AGGGAGGGGA GAACGCGATG GTGAGGGAGG GGAGAACGCG | 120 |
| ACGCGCAGGG GAGGGGGATA ACTCGACGGT GCAGGGAGGT GAGGGGGACG ACGTGACGGC | 180 |
| GCAGGGGAGG GGGGAACCGT CGCGGAAGG GGAAGACCGG GGGGCCGACA AGGTGGTGTT | 240 |
| ACTGGGGTAG GGAGAGGCGG CGTGGAGAAT AGTAACAGAG GGAGGAGTGG TGGTGCTAGG | 300 |
| GTGGAAGAAG GGTAAGAAAG AGGAAGAAAG AGAATTAACA TTATCTTAAC CAAACACCAC | 360 |
| TCTAAATCTA AGGGTTTTCT TTTCCTTTCC TCTCCTCTCC CTTTCTTGAT TCCATTCCCT | 420 |
| TTACCCCGTT GCAACCAAAC GCCCCCTTAT TATGGACCGG AGGAAGTATG TAGAGATGGT | 480 |
| CACAAACTA CTTAAGCTGG TAACTTATAA ATATACTGGG TATTAAATGA ATTAAGTGGC | 540 |
| CACAAATGA CTATAAATTA CTTCGTAATC TTTAGGAACT ATGTTGGTCA CGAAATAACA | 600 |
| TAAACTGGT TATTTAATGG CTTTATGTAG GTACTGCATT CATAAATATA TTTCTAACAT | 660 |
| AATCGTGGTA TGTAGGTGTT TTATAACACA AGGATTAGGT TTACACCAAT GTCATTTTCA | 720 |
| TTAGAATGTA GTTAGAATCA CTTTGGAAC TTTGAAGAGTG ATGACACATT TTTATTATGC | 780 |
| TTTTATGAAA TGTCTTTGTG GTTTTATGA TAGTATTGAG TTTAAGGCAA GTTGGAAGTA | 840 |
| TATGATGGAG AAGTACAGTA TATAGGTGAC AATTGGTTTG CTTGTTTCTA TGAGTTGAAA | 900 |
| GATAAGTAGT ACACGACACT GAGCAATGAC CTCTTCTTAG TTGTAATTTT GTCTTCTCGA | 960 |

| | | | | | | |
|------------|-------------|-------------|-------------|-------------|-------------|------|
| CGTAGTGAAA | GTACAAACAA | GATTATGGCT | TTCAAGCTTC | CAAGATAACG | AGATTGTATG | 1020 |
| AATTTTGTGG | TGTATTTTAC | ATCATTGTTT | TACGTTGGAG | ACAAACTAAA | ACCAATGATG | 1080 |
| AGTTTGTGGA | TTGAGATT | GCCCCTAAGT | CTTATTTACC | CATGGCAAGC | ATGCTGAAAC | 1140 |
| ATGTTAGTCA | AACTTACACA | GCTACAATGT | TTAGGGATT | TGAGCAAAAA | ATTTGGGTAT | 1200 |
| TCTTTGGGTA | CCATTATGTG | AGTTGTTGAC | TATGGATTAA | ACAAAAATCAC | TATATAAAGT | 1260 |
| CTGGAATGAG | AAGCATCCGC | AATTGACACA | CCATGTTACT | TTGATTGTTT | CAACAAGTTT | 1320 |
| ATTAGATGTA | TTTGTAGGAA | TTTTGAAGAG | GCGGAGATGT | TGTGTTATAA | TTGCTTTGGG | 1380 |
| GGTGCTTCAC | ATGCACTCTG | TTAGTGAGAC | ATCTTCAGCT | TATATTTTAA | GGCGGTTAGT | 1440 |
| GAGTATGATT | TTTTTTTTTC | AACTTTTTCG | ATTTCCATGT | AATTAAAAAA | GGTGTGTTGAT | 1500 |
| AAATACATGT | TAAGATAGCC | AAGAAAAGGC | AACTTTCAAA | CAAATAAAAA | AAATTAAGTC | 1560 |
| GCTTAATCAT | TTTTCCAAGT | ACTTTTTTACT | TTTAACACCA | CTTATTACTG | AATCTATAGC | 1620 |
| CGTTAAGAAT | GCATTTTTCAC | GCTCATACAT | GCAAAATCAAG | AACCTCCTCA | TTGAAGGAGA | 1680 |
| TAATTTAGTC | CTCATAAACC | CCGTTAAAGA | CATTTTTCAGC | ATCCAGAGAA | ATTTTCGATTC | 1740 |
| AGTTAAAATT | GCATATATAA | CCAGAGAAAC | AAATTCAGAT | GTTAGTCAGT | CCAGCTACAT | 1800 |
| AGGTCAATGC | CTGAGAGTTT | AAAAGAATCC | GTATCCTTAA | GCATAAGTAG | GTATTGAGGT | 1860 |
| GAGTTACAAA | GGTAAGTTAC | CGGTTACGCA | CCACCTCCAC | CAAACAAGTA | TGGTTAGAAG | 1920 |
| ATACATGTAA | TCGTTTATTT | AGAGTACTAT | TTATAAAAAA | CTTTTAACT | AGAAACAGTT | 1980 |
| GTTTCATTTT | GATATAAGGT | TAATTAGAAT | TCCCGAGCAA | GCAAGAAGGG | GATATAGAGG | 2040 |
| ATAAGGAGGG | CGAGAGAGCG | AGAGAGAGAT | GAAATCAATG | GCGTTATCAA | ACTGCATTCC | 2100 |
| ACAGACACAG | TGCATGCCAT | TGCACAGCAG | CGGGCATTAC | AGGGGCAATT | GATCATGTT | 2160 |
| GTCAATTCCA | TGTAGTTTAA | TTGGAAGACG | AGGTTATTAT | TCACATAAGA | AGAGGAGGAT | 2220 |
| GAGCATGAGT | TGCAGCACAA | GCTCAGGCTC | AAAGTCAGCG | GTTAAAGAAG | CAGGATCAGG | 2280 |
| ATCAGGATCA | GGAGCAGGAG | GATTGCTAGA | CTGCGTAATC | GTTGGAGGTG | GAATTAGCGG | 2340 |
| GCTTTGCATC | GCGCAGGCTC | TTGTACAAA | ACAGTCCTCT | TTATCCCCAA | ATTTTATAGT | 2400 |
| GACAGAGGCC | AAAGACAGAG | TTGGCGGCAA | CATCGTCACT | GTGGAGGCCG | ATGGCTATAT | 2460 |
| CTGGGAGGAG | GGACCCAATA | GCTTCCAGCC | TTCCGACGCG | GTGCTACCA | TGGCGGTAAT | 2520 |
| TCTGTCTCTT | CATTATTCAT | AATCATAATT | CAATTCAATT | CAATTCCTAA | CGTGGAATGT | 2580 |
| GGAATGTGGC | ATGTGCGTAG | GTCGAC | | | | 2606 |

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Pclp_P1a - plastid *clpP* gene promoter top strand PCR primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 4..9
- (D) OTHER INFORMATION: /note= "EcoRI restriction site"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GCGGAATTCA TACTTATTTA TCATTAGAAA G

31

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Pclp_P1b - plastid *clpP* gene promoter bottom strand PCR primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 4..9
- (D) OTHER INFORMATION: /note= "XbaI restriction site"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GCGTCTAGAA AGAACTAAAT ACTATATTTC AC

32

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Pclp_P2b - plastid *clpP*

gene promoter bottom strand PCR primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION: 4..9

(D) OTHER INFORMATION: /note= "NcoI restriction site"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GCGCCATGGT AAATGAAAGA AAGAACTAAA

30

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Trps16_P1a - plastid rps16 gene 3' untranslated region XbaI/HindIII top strand PCR primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION: 4..9

(D) OTHER INFORMATION: /note= "XbaI restriction site"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GCGTCTAGAT CAACCGAAAT TCAATTAAGG

30

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Trps16_p1b - plastid rps16 gene 3' untranslated region XbaI/HindIII bottom strand PCR primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 4..9
- (D) OTHER INFORMATION: /note= "HindIII restriction site"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGCAAGCTTC AATGGAAGCA ATGATAA

27

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "minpsb_U - plastid psbA gene 5' untranslated region 38 nt (blunt/NcoI) including ATG start codon, top strand primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GGGAGTCCCT GATGATTAAA TAAACCAAGA TTTTAC

36

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "minpsb_L - plastid psbA gene 5' untranslated region 38 nt (blunt/NcoI) including ATG start codon (bottom strand primer)"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CATGGTAAAA TCTTGTTTAA TTTAATCATC AGGGACTCCC

40

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "APRTXPl a - top strand PCR primer for amplifying the 5' portion of the mutant Arabidopsis protox gene"
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION: 5..10
 - (D) OTHER INFORMATION: /note= "NcoI restriction site/ATG start codon"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGGACCATGG ATTGTGTGAT TGTCGGCGGA GG

32

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "APRTXPl b - bottom strand PCR primer for amplifying the 5' portion of the mutant Arabidopsis protox gene"
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CTCCGCTCTC CAGCTTAGTG ATAC

24

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 633 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: sugar cane

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..308
- (D) OTHER INFORMATION: /product= "Sugar cane protox-1 partial coding sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```

TTTCCAAGAC TGAAAGTGAG CTGGTAGAAG CAGTTGACCG TGACCTCCGG AAAATGCTTA      60
TAAATCCTAC AGCAGTGGAC CCTTTAGTCC TTGGTGTCCG AGTTTGGCCA CAAGCCATAC      120
CTCAGTTCCT GGTAGGACAT CTTGATCTTC TGGAGGCCGC AAAATCTGCC CTGGACCGAG      180
GTGGCTACGA TGGGCTGTTT CTAGGAGGGA ACTATGTTGC AGGAGTTGCC CTAGGCAGAT      240
GCGTTGAGGG CGCGTATGAG AGTGCCTCGC AAATATATGA CTTCTTGACC AAGTATGCCT      300
ACAAGTGATG AAAGAAGTGG AGTGCTGCTT GTTAATTGTT ATGTTGCATA GATGAGGTGA      360
GACCAGGAGT AGTAAAAGCG TTACGAGTAT TTTTCATTCT TATTTTGTAA ATTGCACTTC      420
TGGTTTTTTC CTGTCAGTAA TTAGTTAGAT TTTAGTTCCT TAGGAGATTG TTCTGTTTAC      480
TGCCCTACAA AAGAATTTTT ATTTTGCATT CGTTTATGAG AGCTGTGCAG ACTTATGTAG      540
CGTTTTTCTG TAAGTACCAA CAAAATCAAA TACTATTCTG TAAGAGCTAA CAGAATGTGC      600
AACTGAGATT GCCTTGGATG AAAAAAAAAA AAA                                     633

```

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

```

Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg
1           5           10           15
Lys Met Leu Ile Asn Pro Thr Ala Val Asp Pro Leu Val Leu Gly Val
20          25          30
Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp
35          40          45
Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Arg Gly Gly Tyr Asp Gly
50          55          60
Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys

```

```

65              70              75              80
Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Tyr Asp Phe Leu Thr
                        85              90              95
Lys Tyr Ala Tyr Lys
              100

```